

DR N-PSDB; AAQ26421


```

Db      121 CTCIDGAVGCIPLCQELSLPMLGCPNRLVAVSGCCCEWVCDSDSIKSDIDDDLLG 180
QY      181 LDASEVELTRNNELIAIGKSSLKRLPVFGTEPRVLFNPLHAHGOKCIYQVTSWQCSKS 240
Db      181 LDASEVELTRNNELIAIGKSSLKRLPVFGTEPRVLFNPLHAHGOKCIYQVTSWQCSKS 240
QY      241 CGTGISTRTVNDNPECRVLKTRICEVRPCGQPVYSSLKKKKCKTKKSPPEPVFTYAG 300
Db      241 CGTGISTRTVNDNPECRVLKTRICEVRPCGQPVYSSLKKKKCKTKKSPPEPVFTYAG 300
QY      301 CSSVKKYRPKYCGSCVDGRCCPTLQTRIVKMFRCEDGEMFSKNMMIOSCKCNCNCPHP 360
Db      301 CSSVKKYRPKYCGSCVDGRCCPTLQTRIVKMFRCEDGEMFSKNMMIOSCKCNCNCPHP 360
QY      361 NEASFRLYSLENDIHKFRD 379
Db      361 NEASFRLYSLENDIHKFRD 379

```

RESULT 3

```

ID      ABB09201 standard; Protein; 379 AA.
XX      ABB09201;
XX      08-JUL-2002 (first entry)
DE      Mouse cyr6 CNN family protein sequence SEQ ID NO:11.
XX      Human; small CCN-like growth factor; SCGF; vulnery; osteopathic;
KW      gene therapy; muscle wasting disease; osteoporosis; wound healing;
KW      tissue regeneration; angiogenesis.
XX      Mus sp.
XX      US2002049304-A1.
XX      25-APR-2002.
XX      14-MAY-2001; 2001US-0853625.
XX      06-JUN-1995; 95US-0468847.
XX      01-APR-1998; 98US-0053587.
XX      (HAST/) HASTINGS G A.
XX      (ADAM/) ADAMS M D.
XX      Hastings GA, Adams MD;
XX      WPI: 2002-382150/41.
XX      Novel isolated polynucleotide sequence encoding a human small CCN-like
XX      growth factor, useful for treating muscle wasting disease, and
XX      osteoporosis -
XX      Disclosure: Fig 2A-D; 33pp; English.
XX      The present invention describes human small CCN-like growth factor
XX      (SCGF). SCGF has vulnery and osteopathic activities, and can be used
XX      in gene therapy. The SCGF polypeptides and polynucleotides can be used
XX      for treating muscle wasting diseases, and osteoporosis, and to stimulate
XX      wound healing and tissue regeneration, to promote angiogenesis and to
XX      stimulate proliferation of vascular smooth muscle and endothelial cell
XX      production. The present sequence represents a CNN family protein which
XX      is given in comparison with the human SCGF in the exemplification of the
XX      present invention.
XX      Sequence 379 AA:

```

```

Query Match      100.0%; Score 2103; DB 23; Length 379;
Best Local Similarity 100.0%; Pred. No. 3,7e-156;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 MSSNFRILAVAVTLLHLTRLALSTCPACACHPLEAPKCAPGVGLVRDGGCKYCAKOL 60
Db      1 MSSNFRILAVAVTLLHLTRLALSTCPACACHPLEAPKCAPGVGLVRDGGCKYCAKOL 60
QY      61 NEDCSKTPQCDHTKLECNFGASSYALGICRAOSEGRPCENSRITYONGESFQPNCKHQ 120
Db      61 NEDCSKTPQCDHTKLECNFGASSYALGICRAOSEGRPCENSRITYONGESFQPNCKHQ 120
QY      121 CTCIDGAVGCIPLCQELSLPMLGCPNRLVAVSGCCCEWVCDSDSIKSDIDDDLLG 180
Db      121 CTCIDGAVGCIPLCQELSLPMLGCPNRLVAVSGCCCEWVCDSDSIKSDIDDDLLG 180
QY      181 LDASEVELTRNNELIAIGKSSLKRLPVFGTEPRVLFNPLHAHGOKCIYQVTSWQCSKS 240
Db      181 LDASEVELTRNNELIAIGKSSLKRLPVFGTEPRVLFNPLHAHGOKCIYQVTSWQCSKS 240
QY      241 CGTGISTRTVNDNPECRVLKTRICEVRPCGQPVYSSLKKKKCKTKKSPPEPVFTYAG 300
Db      241 CGTGISTRTVNDNPECRVLKTRICEVRPCGQPVYSSLKKKKCKTKKSPPEPVFTYAG 300
QY      301 CSSVKKYRPKYCGSCVDGRCCPTLQTRIVKMFRCEDGEMFSKNMMIOSCKCNCNCPHP 360
Db      301 CSSVKKYRPKYCGSCVDGRCCPTLQTRIVKMFRCEDGEMFSKNMMIOSCKCNCNCPHP 360
QY      361 NEASFRLYSLENDIHKFRD 379
Db      361 NEASFRLYSLENDIHKFRD 379

```

RESULT 4

```

ID      AAM35957 standard; Protein; 381 AA.
XX      AAM35957;
XX      05-MAR-1998 (first entry)
DE      Human monocyte mature differentiation factor.
XX      Human; monocyte; mature; differentiation factor; MMDf; macrophage;
KW      cancer; immune activator; tissue culture; infectious disease.
XX      Homo sapiens.
XX      JP09234079-A.
XX      09-SEP-1997.
XX      04-MAR-1996; 96JP-0075236.
XX      04-MAR-1996; 96JP-0075236.
XX      (TOYM ) TOYOBO KK.
XX      WPI: 1997-497320/46.
XX      N-PSDB; AAT97142.
XX      A monocyte mature differentiation factor - useful for the long term
XX      tissue culture of macrophage(s)
XX      Claim 9; Page 12-13; 22pp; Japanese.
XX      The present sequence represents a monocyte mature differentiation
XX      factor (MMDf) which maintains the life of macrophages for long periods
XX      in liquid culture. MMDf can be used as an anti-cancer agent, an immune
XX      activator and to treat infectious diseases.
XX      Sequence 381 AA:

```

```

Query Match      92.2%; Score 1938; DB 18; Length 381;
Best Local Similarity 91.4%; Pred. No. 2.9e-143;
Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

```

QY 1 MSSTFRTLAVALVTLHLTRALSTCPAACHPLEARKCAPGVGLVRDGGCCCKVCAKOL 60
 DB 1 MSSTFRTLAVALVTLHLTRALSTCPAACHPLEARKCAPGVGLVRDGGCCCKVCAKOL 60
 QY 61 NEDSKTOPCDHTGELCNFGASSTALKGICRAOSEGRPCENSRITONGESFOPNCKHQ 120
 DB 61 NEDSKTOPCDHTGELCNFGASSTALKGICRAOSEGRPCENSRITONGESFOPNCKHQ 120
 QY 121 CTCIDGAVGCIPLPQELSLPNLGCPNRLVKGSGCCCEWVCEDESIKDLDDODDL- 178
 DB 121 CTCIDGAVGCIPLPQELSLPNLGCPNRLVKGSGCCCEWVCEDESIKDLDDODDLG 180
 QY 179 --IGLDASEVELTRNNELIAGKSSSLKRLPVFGTEPRVLEPNPLAHGOKCIYOTTSMSQ 236
 DB 181 KELGFDASEVELTRNNELIAGKSSSLKRLPVFGMEPRILYNPL--OGOKCIYOTTSMSQ 238
 QY 237 CSKCGIGISTRTVNDNPECHLVKETRICEVRPGGQVYSSLLKKGKCSKTKSPPEVRF 296
 DB 239 CSKCGIGISTRTVNDNPECHLVKETRICEVRPGGQVYSSLLKKGKCSKTKSPPEVRF 298
 QY 297 TYACSSVKKYRPRYCGSCVDGRCTPLQTRTVKMRFCEDGEMFSKNMMIOSCKCNYN 356
 DB 299 TYACLSVKKYRPRYCGSCVDGRCTPLQTRTVKMRFCEDGEMFSKNMMIOSCKCNYN 358
 QY 357 CPHNEASFRLYSLFNDIHKFRD 379
 DB 359 CPHNEASFRLYSLFNDIHKFRD 381

RESULT 5

AAB90773 standard; Protein; 381 AA.

AAB90773;

15-JUN-2001 (first entry)

Human shear stress-response protein SEQ ID NO: 46.

Human shear stress-response protein; vascular disease;

arteriosclerosis.

Homo sapiens.

MO200125427-A1.

12-APR-2001.

02-OCT-2000; 2000WO-JP06840.

01-OCT-1999; 99JP-0280976.

(KYOW) KYOMA HAKKO KOGYO KK.

(NOJIMA) NOJIMA H.

Nojima H, Yoshisue H, Obayashi M, Oka T, Kawabata A, Sakurada K;

Kuga T, Sekine S, Nakamura Y, Sugano S;

WPI: 2001-266308/27.

N-PSDB: AAH02896.

DNA sequences, proteins encoded by them and antibodies against them

useful in diagnosis and treatment of vascular disease caused by

arteriosclerosis -

Claim 60; Page 345-346; 678pp; Japanese.

The present invention provides the protein and coding sequences of a

number of human shear stress response proteins. These are useful in the

diagnosis, treatment and screening of vascular diseases caused by

arteriosclerosis, including heart failure, post-PTCA restenosis and

hypertension.

XX SQ Sequence 381 AA;
 Query Match 92.2%; Score 1938; DB 22; Length 381;
 Best Local Similarity 91.4%; Pred. No. 2.9e-143;
 Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;
 QY 1 MSSTFRTLAVALVTLHLTRALSTCPAACHPLEARKCAPGVGLVRDGGCCCKVCAKOL 60
 DB 1 MSSTFRTLAVALVTLHLTRALSTCPAACHPLEARKCAPGVGLVRDGGCCCKVCAKOL 60
 QY 61 NEDSKTOPCDHTGELCNFGASSTALKGICRAOSEGRPCENSRITONGESFOPNCKHQ 120
 DB 61 NEDSKTOPCDHTGELCNFGASSTALKGICRAOSEGRPCENSRITONGESFOPNCKHQ 120
 QY 121 CTCIDGAVGCIPLPQELSLPNLGCPNRLVKGSGCCCEWVCEDESIKDLDDODDL- 178
 DB 121 CTCIDGAVGCIPLPQELSLPNLGCPNRLVKGSGCCCEWVCEDESIKDLDDODDLG 180
 QY 179 --IGLDASEVELTRNNELIAGKSSSLKRLPVFGTEPRVLEPNPLAHGOKCIYOTTSMSQ 236
 DB 181 KELGFDASEVELTRNNELIAGKSSSLKRLPVFGMEPRILYNPL--OGOKCIYOTTSMSQ 238
 QY 237 CSKCGIGISTRTVNDNPECHLVKETRICEVRPGGQVYSSLLKKGKCSKTKSPPEVRF 296
 DB 239 CSKCGIGISTRTVNDNPECHLVKETRICEVRPGGQVYSSLLKKGKCSKTKSPPEVRF 298
 QY 297 TYACSSVKKYRPRYCGSCVDGRCTPLQTRTVKMRFCEDGEMFSKNMMIOSCKCNYN 356
 DB 299 TYACLSVKKYRPRYCGSCVDGRCTPLQTRTVKMRFCEDGEMFSKNMMIOSCKCNYN 358
 QY 357 CPHNEASFRLYSLFNDIHKFRD 379
 DB 359 CPHNEASFRLYSLFNDIHKFRD 381

RESULT 6

AAU79761 standard; Protein; 381 AA.

AAU79761;

30-JUL-2002 (first entry)

Human Cyr61 protein.

Human; uterine leiomyoma proliferation; uterine leiomyoma formation;

Cyr61; cytosstatic.

Homo sapiens.

WO200226193-A2.

04-APR-2002.

28-SEP-2001; 2001WO-US30783.

29-SEP-2000; 2000US-236887P.

(AMHP) AMERICAN HOME PROD CORP.

Zhang Z, Sampath D, Zhu Y, Winneker R;

WPI: 2002-383245/41.

N-PSDB: ABK48899.

Preventing uterine leiomyoma formation or inhibiting proliferation of

uterine leiomyoma in subject, comprises modulating or increasing the

level of Cyr61 in leiomyoma tissue -

Disclosure; Fig 6; 92pp; English.

The present invention relates to a method of inhibiting proliferation

QY 1 MSSSTFETLAVATLTLHLTRALSTCPAACHCPLBAPKAPGVGLVBDGGCCCKVCAKOL 60
 DB 1 MSSRIARALAVTLTLHLTRALSTCPAACHCPLBAPKAPGVGLVBDGGCCCKVCAKOL 60
 QY 61 NEDCSKTQPCDHTKGLGECNFGASSSTALKGICRAOSEGRPCVNSRIYONGESFQPNCKHQ 120
 DB 61 NEDCSKTQPCDHTKGLGECNFGASSSTALKGICRAOSEGRPCVNSRIYONGESFQPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNIGCPNPLVKTGGCCCEWVDEDSIKDPMEDQGLLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNIGCPNPLVKTGGCCCEWVDEDSIKDPMEDQGLLG 180
 QY 179 --LGLDASEVELTRNNELIAGKSSSLKRLPVGTGEPVLFNPLHAHGOKCIQVTSWSQ 236
 DB 181 KELGPDASEVELTRNNELIAGKSSSLKRLPVGTGEPVLFNPLHAHGOKCIQVTSWSQ 236
 QY 237 CSKSCGTGISTRTVNDNPECLVKEIRICVRPCGQPVYSSLKKGKCSKTKKSPPEVRF 296
 DB 239 CSKTCGTGISTRTVNDNPECLVKEIRICVRPCGQPVYSSLKKGKCSKTKKSPPEVRF 298
 QY 297 TYAGCSSVKKYRKRYGSCVYDGRCTPQTRVYKMRPCDGEFMSKNVMMIOSCKCNYN 356
 DB 299 TYAGCSSVKKYRKRYGSCVYDGRCTPQTRVYKMRPCDGEFMSKNVMMIOSCKCNYN 358
 QY 357 CPHPNESFRLYSLENDIHKFRD 379
 DB 359 CPHANEAPFPYRLFNDIHKFRD 381

RESULT 8
 ABB05438 standard; Protein: 381 AA.
 ID ABB05438
 AC ABB05438;
 DT 15-APR-2002 (first entry)
 XX Human Cyr61 protein SEQ ID NO:2.
 XX
 XX Human: Cyr61; breast cancer; sex steroid receptor; cytostatic; promoter;
 KM sex steroid response element; cysteine rich heparin-binding protein;
 KM cell proliferation; heparin binding epidermal growth factor;
 KM epidermal growth factor; basic fibroblastic growth factor.
 XX
 OS Homo sapiens.
 PN WO200198359-A2.
 PD 27-DEC-2001.
 XX
 XX 21-JUN-2001; 2001WO-US19823.
 PF 21-JUN-2000; 2000US-213182P.
 PR 16-MAY-2001; 2001US-291510P.
 XX
 PA (AMHP) AMERICAN HOME PROD CORP.
 PI Sampath D, Zhang Z, Winneker R;
 XX
 XX WPI: 2002-147796/19.
 DR N-PSDB; ABA93127, ABA93130.
 XX
 XX Regulation of Cyr61 expression and activity for preventing and
 PT inhibiting breast cancer comprises use of a Cyr61 neutralizing
 PT antibody, an anti-sense oligonucleotide and an antibody which -
 PS Claim 6; Fig 1; 86pp; English.
 XX
 XX The present invention describes a method for the prevention or inhibition
 CC of breast cancer cell proliferation. The method comprises administration
 CC of a compound that inhibits the interaction of a sex steroid receptor
 CC with a sex steroid response element of the Cyr61 (cysteine rich heparin-

CC binding protein) promoter. Cyr61 has cytostatic activity. An antibody (1)
 CC which neutralises Cyr61 can be used to prevent or inhibit breast cancer
 CC cell proliferation by blocking sex steroid induced and growth factor
 CC induced synthesis of Cyr61 DNA, where the growth factor is epidermal.
 CC heparin binding epidermal, or basic fibroblastic growth factor. (1) can be
 CC used to diagnose or stage breast cancer where the level of Cyr61 in a
 CC positive/suspect breast cancer cell is compared to the level in a
 CC normal cell, an increase in the level of Cyr61 compared to the level in
 CC normal tissue indicates the presence of breast cancer. The level of
 CC Cyr61 being determined by exposing the tissues to (1), and an increase
 CC in the level of bound antibody by the suspect/positive cell as compared
 CC to the normal tissue indicates the presence of breast cancer. The present
 CC sequence represents the human Cyr61 protein, which is used in the
 CC exemplification of the present invention.
 CC
 XX
 XX Sequence 381 AA;
 SO
 Query Match 92.28; Score 1938; DB 23; Length 381;
 Best Local Similarity 91.44; Pred. No. 2,9e-143;
 Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;
 QY 1 MSSSTFETLAVATLTLHLTRALSTCPAACHCPLBAPKAPGVGLVBDGGCCCKVCAKOL 60
 DB 1 MSSRIARALAVTLTLHLTRALSTCPAACHCPLBAPKAPGVGLVBDGGCCCKVCAKOL 60
 QY 61 NEDCSKTQPCDHTKGLGECNFGASSSTALKGICRAOSEGRPCVNSRIYONGESFQPNCKHQ 120
 DB 61 NEDCSKTQPCDHTKGLGECNFGASSSTALKGICRAOSEGRPCVNSRIYONGESFQPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNIGCPNPLVKTGGCCCEWVDEDSIKDPMEDQGLLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNIGCPNPLVKTGGCCCEWVDEDSIKDPMEDQGLLG 180
 QY 179 --LGLDASEVELTRNNELIAGKSSSLKRLPVGTGEPVLFNPLHAHGOKCIQVTSWSQ 236
 DB 181 KELGPDASEVELTRNNELIAGKSSSLKRLPVGTGEPVLFNPLHAHGOKCIQVTSWSQ 238
 QY 237 CSKSCGTGISTRTVNDNPECLVKEIRICVRPCGQPVYSSLKKGKCSKTKKSPPEVRF 296
 DB 239 CSKTCGTGISTRTVNDNPECLVKEIRICVRPCGQPVYSSLKKGKCSKTKKSPPEVRF 298
 QY 297 TYAGCSSVKKYRKRYGSCVYDGRCTPQTRVYKMRPCDGEFMSKNVMMIOSCKCNYN 356
 DB 299 TYAGCSSVKKYRKRYGSCVYDGRCTPQTRVYKMRPCDGEFMSKNVMMIOSCKCNYN 358
 QY 357 CPHPNESFRLYSLENDIHKFRD 379
 DB 359 CPHANEAPFPYRLFNDIHKFRD 381

RESULT 9
 AAB43987
 ID AAB43987 standard; Protein: 455 AA.
 AC AAB43987;
 DT 08-FEB-2001 (first entry)
 XX
 XX Human cancer associated protein sequence SEQ ID NO:1432.
 DE
 XX
 XX Human: cancer associated gene; cancer antigen; detection; cancer;
 KM diagnosis; cytostatic; proliferative; vulnerability; immunomodulator;
 KM antidiabetic; antilastmatic; antilethemic; antilethemic; antiviral;
 KM antineoplastic; antineoplastic; antineoplastic; antineoplastic;
 KM dermatological; neuroprotective; thrombolytic; coagulant; nocrotic;
 KM vasotropic; antiproliferative; angiogenic; gene therapy; inflammation;
 KM immune disorder; hematopoietic cell disorder; autoimmune disorder;
 KM allergic reaction; graft versus host disease; organ rejection;
 KM haemostatic; thrombolytic; cardiovascular disorder; infection;
 KM neurological disease; drug screening.
 XX
 OS Homo sapiens.
 XX

[illegible]

Do	433	CPRANEAAPPYRLNDHFRD	455
RESULT 10			
ID	ABG76937		
AC	ABG76937 standard; Protein; 381 AA.		
XX			
XX	ABG76937;		
DT	05-NOV-2002 (first entry)		
DE	Human protein, comprising CYN61, designated SECL.		
XX			
XX	Human; SEC; NOV; immunosuppressive; hepatotropic;		
XX	antifibrotic; angiogenic-associated disorder; diagnostic;		
XX	gene therapy; developmental disorder; immune disease;		
XX	signal transduction pathway disorder; metabolic disorder;		
XX	feeding disorder; obesity; wasting disorder; neurodegenerative disorder;		
XX	Alzheimer's disease; Parkinson's disease; behavioural disorder; allergy;		
XX	asthma; atherosclerosis; cardiomyopathy; angina pectoris;		
XX	autoimmune disease; retinal disease; cirrhosis; diabetes;		
XX	infectious disease; human immunodeficiency virus; HIV; cancer;		
XX	hypertension; hypotension; multiple sclerosis; urinary retention;		
XX	osteoporosis; Crohn's disease; ulcer; neurological disorder; anxiety;		
XX	haemophilia; cirrhosis; immunogen; vaccine.		
OS	Homo sapiens.		
XX			
PN	WO20025705-A2.		
XX			
PD	18-JUL-2002.		
XX			
PE	11-JAN-2002; 2002MO-US00609.		
XX			
PR	11-JAN-2001; 2001US-261013P.		
PR	11-JAN-2001; 2001US-261014P.		
PR	11-JAN-2001; 2001US-261018P.		
PR	11-JAN-2001; 2001US-261026P.		
PR	11-JAN-2001; 2001US-261029P.		
PR	17-AUG-2001; 2001US-313170P.		
PR	10-SEP-2001; 2001US-318410P.		
XX			
XX	(CURA-) CURAGEN CORP.		
PA			
PI	Mezes PS, Rastelli L, Herrmann JL, MacDougall JR, Zhong H;		
PI	Casman SJ, Boldo F, Shinkets RA, Gorman L, Crasta OR, Mysore KK;		
PI	Folkerts O, Martin GB, Eisen A, Spaderna SK, Vernet CM, Bergh C;		
PI	Spytek KA, DiPippo VA, Zernusen BD, Peyman JA, Ellerman K;		
PI	Stone DJ, Grosse WM, Alsdbrook JP, Lepley DM, Rieger DK;		
PI	Burgess CE, Edinger S;		
XX			
DR	WPI: 2002-590675/63.		
DR	N-PSDB; ABS59522.		
XX			
PT	Human SECX/NOVX polypeptide useful for diagnosing, preventing or		
PT	treating disorders associated with aberrant expression or activity of		
PT	SECX/NOVX nucleic acids and proteins e.g., diabetes		
PS			
PS	Example 3; Page 9; 443pp; English.		
XX			
XX	The invention discloses the isolated human polypeptides, and		
XX	polynucleotides encoding them, that have been designated SECX and NOVX.		
XX	The polypeptides can be used for treating, or delaying, the onset of an		
XX	angiogenic-associated disorder or treating a pathological state in a		
XX	subject, preferably a mammal. They can also be used in determining the		
XX	presence of, or predisposition to, a disease associated with altered		
XX	levels of the polypeptides and polynucleotides of any one of the 12		
XX	sequences (SECI-12), for raising antibodies, for identifying an agent		
XX	that binds to, or that modulates the expression or activity of the		
XX	polypeptide, for treating or preventing a NOVX-associated disorder		
XX	(NOVI-8) and as a pharmaceutical composition comprising the polypeptide,		
XX	polynucleotide or an antibody. The polypeptides and polynucleotides are		

CC useful in diagnostic applications where their amounts are assessed, or
 CC for the manufacture of a medicament (e.g. gene therapy) for treating or
 CC preventing disorders or syndromes such as developmental disorders, immune
 CC diseases, signal transduction pathway disorders, metabolic disorders,
 CC feeding disorders (including obesity), wasting disorders,
 CC neurodegenerative disorders (including Alzheimer's disease and
 CC Parkinson's disease), behavioural disorders, allergies, asthma,
 CC atherosclerosis, cardiomyopathy, angina pectoris, autoimmune diseases,
 CC retinal disease, cirrhosis, diabetes, infectious disease (bacterial,
 CC fungal, protozoal and viral e.g. human immunodeficiency virus, HIV),
 CC cancer (e.g. prostate cancer), hypertension, hypotension, multiple
 CC sclerosis, urinary retention, osteoporosis, Crohn's disease, ulcers,
 CC neurological disorders (e.g. anxiety), haemophilia or cirrhosis. They
 CC may also be used as immunogens to produce antibodies specific for the
 CC invention, and as vaccines. Further, they are useful for screening
 CC potential agonist and antagonist compounds. The sequences presented in
 CC ABG76937-ABG76956 are the human SECT-12 and NOV1-8 proteins.

XX
 SQ Sequence 381 AA:

Query Match 91.8%; Score 1930; DB 23; Length 381;
 Best Local Similarity 90.9%; Pred. No. 1.2e-142;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSSSTFRLAVAVTLHLTRALSTCPACCHCPLKAPKAPGVGLVRDGGCCCKVCAKOL 60
 DB 1 MSSRIRALALVVTLLHLTRALSTCPACCHCPLKAPKAPGVGLVRDGGCCCKVCAKOL 60
 QY 61 NEDCSKTOPCDHRTKGLKCNFGASSTALKGICRAOSGRCPEXNSRIYONGESFOPNCKHQ 120
 DB 61 NEDCSKTOPCDHRTKGLKCNFGASSTALKGICRAOSGRCPEXNSRIYONGESFOPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVYKSGCCCEWVCDEDSIKDLDODDL-- 178
 DB 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVYKSGCCCEWVCDEDSIKDLDODDL-- 180
 QY 179 --LGDASEVELTRNNELTAIGKSSLRPLVGTGTEPRVLPFNPJLAHAGOKCIYOTTSMSQ 236
 DB 181 KELGFDASEVELTRNNELTAIGKSSLRPLVGTGTEPRVLPFNPJLAHAGOKCIYOTTSMSQ 238
 QY 237 CSKSCGTGISTRTVNDNPECRILVKEIRICEVRPCGQPVYSLSLKKGKCKSTKSPBVR 296
 DB 239 CSKTCGTGISTRTVNDNPECRILVKEIRICEVRPCGQPVYSLSLKKGKCKSTKSPBVR 298
 QY 297 TYAGCSSVKKYRKRYGSCVDGRCCTPLQTRIVYKMFRCEDGEMFSKNWMIQSKCNTN 356
 DB 299 TYAGCLSVKKYRKRYGSCVDGRCCTPLQTRIVYKMFRCEDGEMFSKNWMIQSKCNTN 358
 QY 357 CPHPNASFRILYSLENDIHKFRD 379
 DB 359 CPHANEAAPFFTRLFNDIHKFRD 381

RESULT 11
 ID AAM35730 standard; Protein; 381 AA.
 XX
 AC AAM35730;
 XX
 DT 27-MAR-1998 (first entry)
 XX
 DE Human cysteine rich protein 61 (Cyr61).
 XX
 KW Cysteine rich protein 61; Cyr61; human;
 KW extracellular matrix signalling molecule; cell adhesion;
 KW cell migration; cell proliferation; angiogenesis; chondrogenesis;
 KW oncogenesis; haematostasis; wound healing; organ regeneration.

XX
 OS Homo sapiens.
 XX
 PN WO9733995-AZ.
 XX
 PD 18-SEP-1997.

XX
 PF 14-MAR-1997; 97WO-US04193.
 XX
 PR 15-MAR-1996; 96US-0013958.
 XX
 PA (MUN1-) MUNIN CORP.
 PI
 XX
 XX lau LF;
 DR WPI: 1997-470875/43.
 DR N-PSDB; AAT94699.
 XX
 XX Isolated and purified cysteine rich protein 61, Cyr61 - useful to
 PT modulate e.g. haematostasis, induce wound healing, promote organ
 PT regeneration etc
 XX
 PS Claim 2; Page 112-113; 133pp; English.

XX
 XX This protein sequence comprises human cysteine rich protein 61
 CC (Cyr61), an extracellular matrix signalling molecule. Its amino
 CC acid sequence was deduced from a human placental cDNA clone (see
 CC AAT94699). Cyr61 polypeptides can be expressed in transformed or
 CC transfected host cells. Cyr61 can be used to modulate
 CC haematostasis, induce wound healing in a tissue, promote organ
 CC regeneration, improve tissue grafting or promote bone or prothesis
 CC implantation (claimed). It can also be used to screen for a
 CC modulator of angiogenesis, chondrogenesis, oncogenesis, cell
 CC adhesion, cell migration, cell proliferation, expand a population
 CC of undifferentiated haematopoietic stem cells in culture and to
 CC screen for a mitogen (claimed). Ex vivo methods for using
 CC mammalian extracellular matrix signalling molecules to prepare
 CC blood products are also provided.

XX
 SQ Sequence 381 AA:

Query Match 91.7%; Score 1929; DB 18; Length 381;
 Best Local Similarity 90.9%; Pred. No. 1.5e-142;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSSSTFRLAVAVTLHLTRALSTCPACCHCPLKAPKAPGVGLVRDGGCCCKVCAKOL 60
 DB 1 MSSRIRALALVVTLLHLTRALSTCPACCHCPLKAPKAPGVGLVRDGGCCCKVCAKOL 60
 QY 61 NEDCSKTOPCDHRTKGLKCNFGASSTALKGICRAOSGRCPEXNSRIYONGESFOPNCKHQ 120
 DB 61 NEDCSKTOPCDHRTKGLKCNFGASSTALKGICRAOSGRCPEXNSRIYONGESFOPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVYKSGCCCEWVCDEDSIKDLDODDL-- 178
 DB 121 CTCIDGAVGCIPLCPQELSLPNLGCNPNRLVYKSGCCCEWVCDEDSIKDLDODDL-- 180
 QY 179 --LGDASEVELTRNNELTAIGKSSLRPLVGTGTEPRVLPFNPJLAHAGOKCIYOTTSMSQ 236
 DB 181 KELGFDASEVELTRNNELTAIGKSSLRPLVGTGTEPRVLPFNPJLAHAGOKCIYOTTSMSQ 238
 QY 237 CSKSCGTGISTRTVNDNPECRILVKEIRICEVRPCGQPVYSLSLKKGKCKSTKSPBVR 296
 DB 239 CSKTCGTGISTRTVNDNPECRILVKEIRICEVRPCGQPVYSLSLKKGKCKSTKSPBVR 298
 QY 297 TYAGCSSVKKYRKRYGSCVDGRCCTPLQTRIVYKMFRCEDGEMFSKNWMIQSKCNTN 356
 DB 299 TYAGCLSVKKYRKRYGSCVDGRCCTPLQTRIVYKMFRCEDGEMFSKNWMIQSKCNTN 358
 QY 357 CPHPNASFRILYSLENDIHKFRD 379
 DB 359 CPHANEAAPFFTRLFNDIHKFRD 381

RESULT 12
 ID AAE05921 standard; Protein; 381 AA.
 XX
 AC AAE05921;
 XX

XX 24-SEP-2001 (first entry)
 XX Human cysteine-rich protein (Cyr61).
 XX Cysteine-rich protein; Cyr61; extracellular matrix signalling molecule;
 XX fibroblast secreted protein; fisp12; connective tissue growth factor;
 XX CTGF; ECM; cell adhesion; cell migration; fibroblast cell proliferation;
 XX angiogenesis; wound healing; integrin receptor; atherosclerosis; tumour;
 XX heart disease; fibrosis; gene therapy; human.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 XX Region 163..229
 XX /note="Cysteine free region"
 XX Domain 212..281
 XX /note="Domain III"
 XX Domain 282..381
 XX /note="Domain IV"
 XX W0200155210-A2.
 XX 02-AUG-2001.
 XX 31-JAN-2001; 2001WO-US03267.
 XX 31-JAN-2000; 2000US-0495448.
 XX 15-MAY-2000; 2000US-0204364.
 XX 06-OCT-2000; 2000US-0238705.
 XX (MUNI-) MUNIN CORP.
 XX Lau LF, Yeung C, Greenspan JA;
 XX WPI: 2001-465561/50.
 XX DR N-PSDB; AAD11221.
 XX Novel human cysteine-rich protein 61 (Cyr61) fragment useful in methods
 XX for screening for modulators of cell adhesion, fibroblast cell
 XX proliferation, angiogenesis and cell migration
 XX Claim 30; Page 171-172; 186pp; English.
 XX The invention relates to extracellular matrix (ECM) signalling
 XX molecules involved in cellular response to growth factors. More
 XX particularly the invention is directed to cysteine-rich protein
 XX (Cyr61), and Cyr61-related proteins such as fibroblast secreted
 XX protein (fisp12) and connective tissue growth factor (CTGF) and
 XX nucleic acid molecules encoding such proteins. The polypeptides
 XX of the invention are members of cysteine-rich secreted protein
 XX family. Human Cyr61 fragment is useful in methods for screening
 XX modulators of cell adhesion, cell migration, fibroblast cell
 XX proliferation, angiogenesis, wound healing and Cyr61-integrin
 XX receptor interaction. Modulator of Cyr61-integrin alphavbeta3
 XX interaction is used for the preparation of a medicament for the
 XX treatment of atherosclerosis, heart disease, tumour metastasis,
 XX fibrosis, tumour growth, disorders associated with inadequate
 XX angiogenesis, aberrant granulation tissue development, aberrant
 XX fibroblast growth and wounds. Polynucleotides of the invention
 XX are useful in gene therapy. The present sequence is human Cyr61
 XX protein.
 XX Sequence 381 AA;
 XX
 QY Query Match 91.7%; Score 1929; DB 22; Length 381;
 DB Best Local Similarity 90.9%; Pred. No. 1.5e-142;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;
 1 MSSSTFTLAVAVTLHLTRALSTCPAACHCPLEAPKAPGVGLVDDGCGCCVCAKOL 60
 1 MSRIARALAVVTLHLTRALSTCPAACHCPLEAPKAPGVGLVDDGCGCCVCAKOL 60

QY 61 NEDCSKTQPCDHTKGLRCNFGASSALKGLICRAOSEGRPCENSRITYONGSEFQPNCKHQ 120
 DB 61 NEDCSKTQPCDHTKGLRCNFGASSALKGLICRAOSEGRPCENSRITYONGSEFQPNCKHQ 120
 QY 121 CTCIDGAVGCTPLCPQELSLPNLGGCPNRLVKKYSGGCCCEWVDCDEDSIKDSLDDQDL-- 178
 DB 121 CTCIDGAVGCTPLCPQELSLPNLGGCPNRLVKKYSGGCCCEWVDCDEDSIKDPMEMDQDLIG 180
 QY 179 --LGLDASVEHLTRNNELIAGKSSLRPLVFETBERVLEPNLHAIGOKCIYQTSMSQ 236
 DB 181 KELGFDASEVELTRNNELIAGKRSRLKRLPVEFMERILLYNPL--OGOKCIYQTSMSQ 238
 QY 237 CSKSCGTGSTRVNDNDPECRVLYKTRICEVRPCGQPVYSSLRKGRKSKTKSPPEVR 296
 DB 239 CSKTCGTGISTRVNDNDPECRVLYKTRICEVRPCGQPVYSSLRKGRKSKTKSPPEVR 298
 QY 297 TYAGCCSVKKYRPYKCGSCVDGRCTPLQRTYVRRRCEDGEMFSKNVMIIQSCKNYN 356
 DB 299 TYAGCLSVKKYRPYKCGSCVDGRCTPLQRTYVRRRCEDGEMFSKNVMIIQSCKNYN 358
 QY 357 CPHPEASFRLYSLFNDIHKFRD 379
 DB 359 CPHANEAFPEYRLFNDIHKFRD 381
 RESULT 13
 ABB09202
 ID ABB09202 standard; Protein: 374 AA.
 XX ABB09202;
 XX 08-JUL-2002 (first entry)
 XX HCGF CNN family protein sequence SEQ ID NO:12.
 XX Human; small CCN-like growth factor; SCGF; vulnary; osteopathic;
 XX gene therapy; muscle wasting disease; osteoporosis; wound healing;
 XX tissue regeneration; angiogenesis.
 XX Unidentified.
 XX US2002049304-A1.
 XX 25-APR-2002.
 XX 14-MAY-2001; 2001US-0853625.
 XX 06-JUN-1995; 95US-0468847.
 XX 01-APR-1998; 98US-0053587.
 XX (HAST/) HASTINGS G A.
 XX (ADAM/) ADAMS M D.
 XX Hastings GA, Adams MD;
 XX WPI: 2002-382150/41.
 XX Novel isolated polynucleotide sequence encoding a human small CCN-like
 XX growth factor, useful for treating muscle wasting disease, and
 XX osteoporosis -
 XX Disclosure; Fig 2A-D; 33pp; English.
 XX The present invention describes human small CCN-like growth factor
 XX (SCGF). SCGF has vulnary and osteopathic activities, and can be used
 XX in gene therapy. The SCGF polypeptides and polynucleotides can be used
 XX for treating muscle wasting diseases, and osteoporosis, and to stimulate
 XX wound healing and tissue regeneration, to promote angiogenesis and to
 XX stimulate proliferation of vascular smooth muscle and endothelial cell
 XX production. The present sequence represents a CNN family protein which
 XX is given in comparison with the human SCGF in the exemplification of the
 XX present invention.

SQ Sequence 374 AA;
 Query Match 86.5%; Score 1819.5; DB 23; Length 374;
 Best Local Similarity 88.3%; Pred. No. 5.1e-134;
 Matches 332; Conservative 10; Mismatches 27; Indels 7; Gaps 3;

QY 1 MSSSTFRTLAVALTLLHTRIALSTCPACCHGLEAPKCAPGVGLVDDGGCCCKVCAKOL 60
 D 1 MSSRIVELALVLTLLHTRVGLSTCPADCHGLEAPKCAPGVGLVDDGGCCCKVCAKOL 60
 QY 61 NEDCSKTQPCDHTKGLGECNFGASTALKGICRAQSEGRCEYNSRIYONGESFOPNCKHQ 120
 D 61 NEDCSKTQPCDHTKGLGECNFGASTALKGICRAQSEGRCEYNSRIYONGESFOPNCKHQ 120
 QY 121 CTCIDGAVG-CIPLCPOEISLPNIGCPNRLVYVSGQCEEWYCDSDSDDDDL- 178
 D 121 CTCIDGAVG-CIPLCPOEISLPNIGCPNRLVYVSGQCEEWYCDSDSDDDDL 180
 QY 179 ---LGLDASEVELTRNNELIAIGSSSLKRLPVFGTEPRVLFNPLAHAGOKCIYQTTSMS 235
 D 181 GKGLGFDASEVELTRNNELIAVGSSSLKRLPVFGEMPRILYNPL--QGCKIYQTTSMS 238
 QY 236 QCSKSCGTGISTRTVNDNPECRILVKEIRICEVRPCQPYVSSLKKGKCKSKTKKSPPEVR 295
 D 239 QCSKSCGTGISTRTVNDNPECRILVKEIRICEVRPCQPYVSSLKKGKCKSKTKKSPPEVR 298
 QY 296 FTYAGGSSVKKYKPKYCGSCVDGRCTPLQTRVYKMRFCEDGEMFSKNVMMIOSSKCN 355
 D 299 FTYAGGSSVKKYKPKYCGSCVDGRCTPLQTRVYKMRFCEDGEMFSKNVMMIOSSKCN 358
 QY 356 NCPHPNEASFRLYSLE 371
 D 359 NCPHPNEASFRLYSLE 374

RESULT 14
 AAR90919
 ID AAR90919 standard; Protein: 375 AA.
 AC AAR90919;
 DT 25-MAR-2003 (updated)
 DT 25-JUN-1996 (first entry)
 XX Connective tissue growth factor-2.
 DE Connective tissue growth factor-2.
 XX
 KM CTGF-2; connective tissue growth factor-2; secreted protein;
 KM cartilaginous growth; skeletal; embryo; cell growth; morphogenesis;
 KM insulin-like growth factor; fibroblast growth factor; Cry61.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..24
 FT Protein /label= signal_peptide
 FT /label= mature_protein
 XX
 PN WO9601896-A1.
 XX
 PD 25-JAN-1996.
 PF 12-JUL-1994; 94MO-US07736.
 PR 12-JUL-1994; 94MO-US07736.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Adams MD, Li H;
 XX
 DR WPI: 1996-097626/10.
 DR N-PSDB; AAT12653.
 XX

PT Connective tissue growth factor-2 and DNA encoding it - useful to
 enhance the repair of connective and support tissue, and to enhance
 wound healing
 PT
 XX
 XX Claim 1; Fig 1A-C; 46pp; English.
 PS
 CC Connective tissue growth factor-2 (CTGF-2) is encoded by a cDNA
 CC (AAT12653) isolated from a human foetal lung cDNA library. The CTGF
 CC polypeptides are structurally and functionally related to a family
 CC of growth factors which include IGF (insulin-like growth factor)
 CC PDGF (platelet-derived growth factor), and FGF (fibroblast growth
 CC factor). CTGF-2 exhibits 89 percent identity and 93 percent similarity
 CC to Cry61. Cry61 is a growth factor-inducible immediate early gene
 CC initially identified in serum-stimulated mouse fibroblasts. It encodes
 CC a member of an emerging family of secreted proteins which are also a
 CC group of cysteine-rich proteins. This group of GFs are important for
 CC normal growth, differentiation, morphogenesis of the cartilaginous
 CC skeleton of an embryo and cell growth.
 CC (updated on 25-MAR-2003 to correct PN field.)
 XX

SQ Sequence 375 AA;
 Query Match 86.5%; Score 1819.5; DB 17; Length 375;
 Best Local Similarity 88.3%; Pred. No. 5.1e-134;
 Matches 332; Conservative 10; Mismatches 27; Indels 7; Gaps 3;

QY 1 MSSSTFRTLAVALTLLHTRIALSTCPACCHGLEAPKCAPGVGLVDDGGCCCKVCAKOL 60
 D 1 MSSRIVELALVLTLLHTRVGLSTCPADCHGLEAPKCAPGVGLVDDGGCCCKVCAKOL 60
 QY 61 NEDCSKTQPCDHTKGLGECNFGASTALKGICRAQSEGRCEYNSRIYONGESFOPNCKHQ 120
 D 61 NEDCSKTQPCDHTKGLGECNFGASTALKGICRAQSEGRCEYNSRIYONGESFOPNCKHQ 120
 QY 121 CTCIDGAVG-CIPLCPOEISLPNIGCPNRLVYVSGQCEEWYCDSDSDDDDL- 178
 D 121 CTCIDGAVG-CIPLCPOEISLPNIGCPNRLVYVSGQCEEWYCDSDSDDDDL 180
 QY 179 ---LGLDASEVELTRNNELIAIGSSSLKRLPVFGTEPRVLFNPLAHAGOKCIYQTTSMS 235
 D 181 GKGLGFDASEVELTRNNELIAVGSSSLKRLPVFGEMPRILYNPL--QGCKIYQTTSMS 238
 QY 236 QCSKSCGTGISTRTVNDNPECRILVKEIRICEVRPCQPYVSSLKKGKCKSKTKKSPPEVR 295
 D 239 QCSKSCGTGISTRTVNDNPECRILVKEIRICEVRPCQPYVSSLKKGKCKSKTKKSPPEVR 298
 QY 296 FTYAGGSSVKKYKPKYCGSCVDGRCTPLQTRVYKMRFCEDGEMFSKNVMMIOSSKCN 355
 D 299 FTYAGGSSVKKYKPKYCGSCVDGRCTPLQTRVYKMRFCEDGEMFSKNVMMIOSSKCN 358
 QY 356 NCPHPNEASFRLYSLE 371
 D 359 NCPHPNEASFRLYSLE 374

RESULT 15
 AAY31620
 ID AAY31620 standard; Protein: 375 AA.
 AC AAY31620;
 DT 02-NOV-1999 (first entry)
 DE Human CTGF-2.
 XX
 KM Connective tissue growth factor-2; CTGF-2; wound healing; bone disorder;
 KM skin disorder; acne; burn; UV damage; stabilisation; tissue implant.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..24
 FT /label= signal_peptide

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:03:52 ; Search time 13.9632 Seconds
(without alignments)
1148.438 Million cell updates/sec

Title: US-09-495-448a-2

Perfect score: 2103
Sequence: 1 MSSSTFRLAAVAVTLHLTR.....PNEASFRLYSIFNDIKFRD 379

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfill1st.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2103	100.0	379	1	US-08-468-847B-11
2	2103	100.0	379	4	US-09-142-569-2
3	1938	92.2	381	4	US-09-348-815-2
4	1929	91.7	381	4	US-09-142-569-4
5	1819.5	86.5	374	1	US-08-468-847B-12
6	1812.5	86.2	375	2	US-08-459-101A-2
7	1646.5	78.3	347	1	US-08-468-847B-13
8	958	45.6	347	4	US-09-582-337-2
9	957.5	45.5	348	4	US-09-292-036-3
10	953.5	45.3	348	1	US-08-468-847B-15
11	953.5	45.3	348	4	US-09-142-569-6
12	950	45.2	349	1	US-08-167-628-2
13	950	45.2	349	1	US-08-386-680-2
14	950	45.2	349	1	US-08-459-717-2
15	950	45.2	349	1	US-08-712-302-2
16	950	45.2	349	2	US-08-880-031-2
17	950	45.2	349	3	US-09-054-368-2
18	950	45.2	349	3	US-09-097-179-2
19	950	45.2	349	3	US-09-054-274-2
20	950	45.2	349	3	US-09-080-715-2
21	950	45.2	349	3	US-09-056-704-2
22	950	45.2	349	4	US-09-292-036-4
23	950	45.2	349	4	US-09-253-316-26
24	950	45.2	349	4	US-09-142-569-8
25	950	45.2	349	4	US-09-461-688-2
26	950	45.2	349	5	PCT-US96-08140-2
27	944	44.9	347	4	US-09-187-478-2

28	944	44.9	347	4	US-09-292-036-2	Sequence 2, Appl
29	941.5	44.8	348	1	US-08-468-847B-14	Sequence 14, Appl
30	861	40.9	351	1	US-08-468-847B-16	Sequence 16, Appl
31	839.5	39.9	357	1	US-08-468-847B-17	Sequence 17, Appl
32	839.5	39.9	357	4	US-09-253-316-25	Sequence 25, Appl
33	773	36.8	345	4	US-09-182-145-3	Sequence 3, Appl
34	773	36.8	345	4	US-09-182-145-6	Sequence 6, Appl
35	773	36.8	367	4	US-09-182-145-4	Sequence 4, Appl
36	773	36.8	367	4	US-09-182-145-8	Sequence 8, Appl
37	772	36.7	345	4	US-09-182-145-5	Sequence 5, Appl
38	772	36.7	345	4	US-09-182-145-21	Sequence 21, Appl
39	772	36.7	367	4	US-09-182-145-7	Sequence 7, Appl
40	772	36.7	367	4	US-09-182-145-22	Sequence 22, Appl
41	758.5	36.1	345	4	US-09-182-145-11	Sequence 11, Appl
42	758.5	36.1	367	4	US-09-182-145-12	Sequence 12, Appl
43	622	29.6	339	4	US-09-182-145-36	Sequence 36, Appl
44	622	29.6	354	4	US-09-182-145-37	Sequence 37, Appl
45	622	29.6	354	4	US-09-253-316-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-468-847B-11
; Sequence 11, Application US/08468847B
; Patent No. 5780263
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
; TITLE OF INVENTION: Human CCN-Like Growth Factor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,847B
; FILING DATE: 6 June 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J. G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-08-468-847B-11
; Query Match 100.0%; Score 2103; DB 1; Length 379;
; Best Local Similarity 100.0%; Pred. No. 4.2e-170;
; Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MSSSTFRLAAVAVTLHLTRALSTCPACACPLEAPKAPGVGLVRDCCCKKVCARQL 60
|||||

Db 1 MSSSTFTLAAVATLTLHLTRALSTCPAACHCPLEAPKAPGVGLVRDGGCCCKVCAKOL 60
QY 61 NEDOSKTOPCDHRTKGLCECNFGASTALKGICRAOSEGRPCPEYNSRIYONGESFOPNCKHQ 120
Db 61 NEDOSKTOPCDHRTKGLCECNFGASTALKGICRAOSEGRPCPEYNSRIYONGESFOPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVKVSGGCCCEWVDEDSIKSLDDODDLG 180
Db 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVKVSGGCCCEWVDEDSIKSLDDODDLG 180
QY 181 LDASEVELTRNNELIAGKSSSLKRLPVFETEPVLFNPLHAHQKCIYOTTSMSCSKS 240
Db 181 LDASEVELTRNNELIAGKSSSLKRLPVFETEPVLFNPLHAHQKCIYOTTSMSCSKS 240
QY 241 CGTISTRTVNDNPECLVETRICVPRCPGQPVYSSLLKGGKCSKTKKSPPEVRYTAA 300
Db 241 CGTISTRTVNDNPECLVETRICVPRCPGQPVYSSLLKGGKCSKTKKSPPEVRYTAA 300
QY 301 CSSVKATRPKYCGSCVDGRCTPLQRTVYKMRFRCEDEGEFSKNVMMIOCKCNVNCPPH 360
Db 301 CSSVKATRPKYCGSCVDGRCTPLQRTVYKMRFRCEDEGEFSKNVMMIOCKCNVNCPPH 360
QY 361 NEASFRLYSLFNDIHKFRD 379
Db 361 NEASFRLYSLFNDIHKFRD 379

RESULT 2

US-09-142-569-2
; Sequence 2, Application US/09142569
; Patent No. 6413735

GENERAL INFORMATION:

APPLICANT: Lau, Lester F.
TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/142,569

FILING DATE: 02-Apr-1999

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Clough, David W.
REGISTRATION NUMBER: 36,107

REFERENCE/DOCKET NUMBER: 28758/33766

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 379 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:

NAME/KEY: misc.feature
OTHER INFORMATION: "Mouse Cyt61 amino acid sequence"

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSSSTFTLAAVATLTLHLTRALSTCPAACHCPLEAPKAPGVGLVRDGGCCCKVCAKOL 60
Db 1 MSSSTFTLAAVATLTLHLTRALSTCPAACHCPLEAPKAPGVGLVRDGGCCCKVCAKOL 60
QY 61 NEDOSKTOPCDHRTKGLCECNFGASTALKGICRAOSEGRPCPEYNSRIYONGESFOPNCKHQ 120
Db 61 NEDOSKTOPCDHRTKGLCECNFGASTALKGICRAOSEGRPCPEYNSRIYONGESFOPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVKVSGGCCCEWVDEDSIKSLDDODDLG 180
Db 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVKVSGGCCCEWVDEDSIKSLDDODDLG 180
QY 181 LDASEVELTRNNELIAGKSSSLKRLPVFETEPVLFNPLHAHQKCIYOTTSMSCSKS 240
Db 181 LDASEVELTRNNELIAGKSSSLKRLPVFETEPVLFNPLHAHQKCIYOTTSMSCSKS 240
QY 241 CGTISTRTVNDNPECLVETRICVPRCPGQPVYSSLLKGGKCSKTKKSPPEVRYTAA 300
Db 241 CGTISTRTVNDNPECLVETRICVPRCPGQPVYSSLLKGGKCSKTKKSPPEVRYTAA 300
QY 301 CSSVKATRPKYCGSCVDGRCTPLQRTVYKMRFRCEDEGEFSKNVMMIOCKCNVNCPPH 360
Db 301 CSSVKATRPKYCGSCVDGRCTPLQRTVYKMRFRCEDEGEFSKNVMMIOCKCNVNCPPH 360
QY 361 NEASFRLYSLFNDIHKFRD 379
Db 361 NEASFRLYSLFNDIHKFRD 379

RESULT 3

US-09-348-815-2
; Sequence 2, Application US/09348815
; Patent No. 6534630

GENERAL INFORMATION:

APPLICANT: LI, HAODONG
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR-2
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:

ADDRESSEE: HUMAN GENOME SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: US
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/348,815

FILING DATE: 08-Jul-1999

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: JONATHAN L. KLEIN
REGISTRATION NUMBER: 41,119

REFERENCE/DOCKET NUMBER: PFI26PID1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-309-8504

TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 381 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:

NAME/KEY: misc.feature
OTHER INFORMATION: "Mouse Cyt61 amino acid sequence"

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Query Match 100.0%; Score 2103; DB 4; Length 379;
Best Local Similarity 100.0%; Pred. No. 4,2e-170;

Query Match 92.2%; Score 1938; DB 4; Length 381;

Best Local Similarity 91.4%; Pred. No. 3.8e-156;
Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2.

QY	1	MSSTFRTLA	VATL	LIHL	TR	LA	LS	PC	AC	HP	LE	AP	RC	AG	VG	LV	DD	GG	CC	VC	AK	OL	60
			:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	
Db	1	MSRIRAKAL	AVAT	LIHL	TR	LA	LS	PC	AC	HP	LE	AP	RC	AG	VG	LV	DD	GG	CC	VC	AK	OL	60
QY	61	NEDCSKT	OP	CD	HT	K	TE	CN	F	GA	SS	T	AL	G	I	RA	O	S	E	G	R	P	Y
			:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	
Db	61	NEDCSKT	OP	CD	HT	K	TE	CN	F	GA	SS	T	AL	G	I	RA	O	S	E	G	R	P	Y
QY	121	CTCIDG	AV	GC	I	P	L	C	P	Q	E	L	S	P	N	L	G	C	P	N	R	L	V
			:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	
Db	121	CTCIDG	AV	GC	I	P	L	C	P	Q	E	L	S	P	N	L	G	C	P	N	R	L	V
QY	179	--	-	L	G	I	D	A	S	V	E	L	T	R	N	N	E	L	I	A	I	G	K
			:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	
Db	181	KEIG	F	D	A	S	V	E	L	T	R	N	N	E	L	I	A	I	G	K	S	L	R
QY	237	CSKSC	T	G	I	S	T	E	N	T	D	N	E	C	R	L	Y	K	T	R	I	C	E
			:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	
Db	239	CSKTG	T	G	I	S	T	E	N	T	D	N	E	C	R	L	Y	K	T	R	I	C	E
QY	297	TYAG	C	S	V	A	K	K	P	K	Y	C	G	S	C	V	D	R	C	T	P	O	L
			:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	
Db	299	TYAG	C	S	V	A	K	K	P	K	Y	C	G	S	C	V	D	R	C	T	P	O	L
QY	357	CPHP	N	E	A	S	F	L	S	T	E	N	D	I	H	K	F	R	D	I	H	K	F
			:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	
Db	359	CPHP	N	E	A	S	F	L	S	T	E	N	D	I	H	K	F	R	D	I	H	K	F

RESULT 4

US-09-142-569-4
; Sequence 4, Application US/09142569
; Patent No. 6413735

GENERAL INFORMATION:

APPLICANT: Lau, Lester F.
TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
NUMBER OF DRAWINGS: 17

NUMBER OF SEQUENCES: 1 /
CORRESPONDENCE ADDRESS:

ADDRESS: Marshall, O'Boole, Ger
STREET: 6300 Sears Tower, 233 Sou
CITY: Chicago

CITY: Chicago
STATE: Illinois
COUNTRY: U.S.A.

COUNTRY: United States of America
ZIP: 60606-6402

```

;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; CONVERTED: TIV 02 00000400

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
COMMAND: Batchfile: Del.bat #10

```

; SOFTWARE: PatentIn Release #1.0,
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: 75/004143 E60

```

APPLICATION NUMBER: 05/09/142,509
FILING DATE: 02-Apr-1999
CLASSIFICATION: H04B1/00

SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-142-569-4

Query Match	Score	DB 4	Length
91.7%	1929	DB 4	381

Desc	DocId	Gene	Start	End	Strand	Score	Match	Conservative	Mismatches	Indels	Gaps
Matches	348;							10;			
Qy	1	MSSTFRTLAIVAVTLLHLTLRLALSTPCPACHCPLFAPKCAPGVGLVBDGGCCCKVCAKOL	60								
Db	1	MSRIRALALAVVTLHLHLTRLALSTPCPACHCPLFAPKCAPGVGLVBDGGCCCKVCAKOL	60								
Qy	61	NEDCSKTQPCDHTKGLGECNFGASSTALGICRAOSEGNPCYSRIYQNESEFPQNC HQ	120								
Db	61	NEDCSKTQPCDHTKGLGECNFGASSTALGICRAOSEGNPCYSRIYQNESEFPQNC HQ	120								
Qy	121	CTCICIDAVGCIPLCQDELPLNLGCPNRLVYKVSQCCCEWVYCDSDSITKSLDDQDL	178								
Db	121	CTCICIDAVGCIPLCQDELPLNLGCPNRLVYKVSQCCCEWVYCDSDSITKSLDDQDL	180								
Qy	179	--LGIDASEVELTRNNELIAGKSSLRPLVEFGETEPRVLEPNPLAHGQKCIVQYTTSMQ	236								
Db	181	KELGFDASEVELTRNNELIAGKSGRLKRLPVFGEMERILYNP--GQKCIQYTTSMQ	238								
Qy	237	CSKSGCGTISTRYTNDNPECRLYKETRIRICEVAPCGQPYSSLKKGKKCSKTKSPEVYR	296								
Db	239	CSKTCTGTISTRYTNDNPECRLYKETRIRICEVAPCGOPYSSLKKGKKCSKTKSPEVYR	298								
Qy	297	TYAGCSYKVKYRPKVCGSCVDRGCTPLQTFYVKKRPFCEGEGEMSKNNMMIIOCKCNYN	356								
Db	299	TYAGCLSYKVKYRPKVCGSCVDRGCTPLQTFYVKKRPFCEGEGESKNNMMIIOCKCNYN	358								
Qy	357	CPHPNEASPRLYSLFNDIHKFRD	379								
Db	359	CPHPNEAEPFRLFNDIHKFRD	381								

RESULT 5

US-08-468-847B-12
; Sequence 12, Application US/08468847E
; Patent No. 5780263

```

; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.

```

```

; TITLE OF INVENTION: Human CCN-Like Growth Fa
;
; NUMBER OF SEQUENCES: 20

```

```

;
;  CORRESPONDENCE ADDRESS:
;  ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLIAN
;

```

ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD

```
; CITY: ROSELAND
; STATE: NEW JERSEY
;
```

```

;      COUNTRY:  USA
;      ZIP:      07068
;

```

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
;

```

```

;
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
;

```

LENGTH: 374 AMINO ACIDS
 TYPE: AMINO ACID
 STRANDEDNESS:
 TOPOLOGY: LINEAR
 MOLECULE TYPE: PROTEIN
 US-08-468-847B-12

Query Match 86.5%; Score 1819.5; DB 1; Length 374;
 Best Local Similarity 88.3%; Pred. No. 3,9e-146;
 Matches 332; Conservative 10; Mismatches 27; Indels 7; Gaps 3;

QY 1 MSSTRTTAVANTLTLTLALSTCPACCHCLEAPKAPGVGLVRDGGCCCKVCAKOL 60
 DB 1 MSSRIARELAVTLTLTLTRVGLSTCPADCHCLEAPKAPGVGLVRDGGCCCKVCAKOL 60
 QY 61 NECCSKTOPCDHTKGLKGCNFGASTALKGICRAOSEGRPEYNSRIYONGESFOPNCKHQ 120
 DB 61 NECCSKTOPCDHTKGLKGCNFGASTALKGICRAOSEGRPEYNSRIYONGESFOPNCKHQ 120
 QY 121 CTCIDGAVG-CIPLCQDELSPNLGCPNRLVYVSGCCCEWYCDSDSIKSDDDDDL- 178
 DB 121 CTCIGMRGACIPLCQDELSPNLGCPNRLVYVSGCCCEWYCDSDSIKSDDDDDL- 180
 QY 179 ---LGIDASEVELTRNNELIAIGKSSLRKLPVFGTEPRVLFNPLAHAGOKCIYOTTSS 235
 DB 181 GKGLGDASEVELTRNNELIAVKGSSLRKLPVFGTEPRVLFNPLAHAGOKCIYOTTSS 238
 QY 236 QCSKSGGTGISTRTVNDNPECRVLYKETRICVEVRPCGOPYSSSLKKGKSKTKKSPPEVR 295
 DB 239 QCSKSGGTGISTRTVNDNPECRVLYKETRICVEVRPCGOPYSSSLKKGKSKTKKSPPEVR 298
 QY 296 FTYAGSSVYKRYKPKYCGSCVDGRCTPLOTRTVKNRPFCEDEGEMSKNMIOSKCNV 355
 DB 299 FTYAGCLSVYKRYKPKYCGSCVDGRCTPLOTRTVKNRPFCEDEGEMSKNMIOSKCNV 358
 QY 356 NCPHNEASPRLSLF 371
 DB 359 NCPHNEAAPPFYRLF 374

RESULT 6

US-08-459-101A-2
 Sequence 2, Application US/08459101A

Patent No. 5945300
 GENERAL INFORMATION:
 APPLICANT: LI, ET AL.
 TITLE OF INVENTION: Connective Tissue Growth Factor-2
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
 ADDRESSEE: CECCHI, STEWART & OLSTEIN
 STREET: 6 BECKER FARM ROAD
 CITY: ROSELAND
 STATE: NEW JERSEY
 COUNTRY: USA
 ZIP: 07068
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 INCH DISKETTE
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: WORD PERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/459,101A
 FILING DATE: June 2, 1995
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/07736
 FILING DATE: 12 JUL 94
 ATTORNEY/AGENT INFORMATION:
 NAME: FERRARO, GREGORY D.
 REGISTRATION NUMBER: 36,134
 REFERENCE/DOCKET NUMBER: 325800-317
 TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744
 INFORMATION FOR SEQ ID NO:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 375 AMINO ACIDS
 TYPE: AMINO ACID
 STRANDEDNESS:
 TOPOLOGY: LINEAR
 MOLECULE TYPE: PROTEIN
 US-08-459-101A-2

Query Match 86.2%; Score 1812.5; DB 2; Length 375;
 Best Local Similarity 88.0%; Pred. No. 1,5e-145;
 Matches 331; Conservative 10; Mismatches 28; Indels 7; Gaps 3;

QY 1 MSSTRTTAVANTLTLTLALSTCPACCHCLEAPKAPGVGLVRDGGCCCKVCAKOL 60
 DB 1 MSSRIARELAVTLTLTLTRVGLSTCPADCHCLEAPKAPGVGLVRDGGCCCKVCAKOL 60
 QY 61 NECCSKTOPCDHTKGLKGCNFGASTALKGICRAOSEGRPEYNSRIYONGESFOPNCKHQ 120
 DB 61 NECCSKTOPCDHTKGLKGCNFGASTALKGICRAOSEGRPEYNSRIYONGESFOPNCKHQ 120
 QY 121 CTCIDGAVG-CIPLCQDELSPNLGCPNRLVYVSGCCCEWYCDSDSIKSDDDDDL- 178
 DB 121 CTCIGMRGACIPLCQDELSPNLGCPNRLVYVSGCCCEWYCDSDSIKSDDDDDL- 180
 QY 179 ---LGIDASEVELTRNNELIAIGKSSLRKLPVFGTEPRVLFNPLAHAGOKCIYOTTSS 235
 DB 181 GKGLGDASEVELTRNNELIAVKGSSLRKLPVFGTEPRVLFNPLAHAGOKCIYOTTSS 238
 QY 236 QCSKSGGTGISTRTVNDNPECRVLYKETRICVEVRPCGOPYSSSLKKGKSKTKKSPPEVR 295
 DB 239 QCSKSGGTGISTRTVNDNPECRVLYKETRICVEVRPCGOPYSSSLKKGKSKTKKSPPEVR 298
 QY 296 FTYAGSSVYKRYKPKYCGSCVDGRCTPLOTRTVKNRPFCEDEGEMSKNMIOSKCNV 355
 DB 299 FTYAGCLSVYKRYKPKYCGSCVDGRCTPLOTRTVKNRPFCEDEGEMSKNMIOSKCNV 358
 QY 356 NCPHNEASPRLSLF 371
 DB 359 NCPHNEAAPPFYRLF 374

RESULT 7

US-08-468-847B-13
 Sequence 13, Application US/08468847B

Patent No. 5780263
 GENERAL INFORMATION:
 APPLICANT: Hastings, Gregg A. and Adams, Mark D.
 TITLE OF INVENTION: Human CCN-Like Growth Factor
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
 ADDRESSEE: CECCHI, STEWART & OLSTEIN
 STREET: 6 BECKER FARM ROAD
 CITY: ROSELAND
 STATE: NEW JERSEY
 COUNTRY: USA
 ZIP: 07068
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 INCH DISKETTE
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: WORD PERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/468,847B
 FILING DATE: 6 June 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:

RESULT 8	
US-09-582-337-2	
Sequence 2, Application US/09582337	
Patent No. 6562618	
GENERAL INFORMATION:	
APPLICANT: Japan Tobacco, Inc.	
TITLE OF INVENTION: Monoclonal Antibody Against Connective Tissue Growth Factor	
TITLE OF INVENTION: and Medicinal Uses Thereof	
FILE REFERENCE: JI-009PCT	
CURRENT APPLICATION NUMBER: US/09/582,337	
CURRENT FILING DATE: 2000-06-23	
PRIOR APPLICATION NUMBER: JP P1997-367699	
PRIOR FILING DATE: 1997-12-25	
PRIOR APPLICATION NUMBER: JP P1998-356183	
PRIOR FILING DATE: 1998-12-15	
NUMBER OF SEQ ID NOS: 27	
SOFTWARE: PatentIn Ver. 2.0	
SEQ ID NO 2	
LENGTH: 347	
TYPE: PRT	
ORGANISM: Rat	
US-09-582-337-2	
Query Match	45.6%; Score 958; DB 4;
Best Local Similarity	46.4%; Pred. No. 2,66-73; Length 347;

[illegible]

QY 178 LGLDASEVELTRNNELIAGKSSILKRL-PVGTETBRVLFNPLHAHGQKCIYOTMSQ 236
Db 169 -----KDRTAAGPALAAVRLDTEGPDPTM-----RANCLVOTEMSA 207
QY 237 CSKSCGTGISTRTVNDPNDPCEKRLVETRICVPRGQGVYSSILKKGKCSKTKSPPEVRF 296
Db 208 CSKTCGAGISTRTVNDPNDPCEKRLVETRICVPRGQGVYSSILKKGKCSKTKSPPEVRF 267
QY 297 TYAGCSSLVKKRYPRKYGSCVDGRCCPTLQRTVYKMRFCEDGEMFSKNVMIOSCCN 356
Db 268 ELSCGTSVKTYRAKFCGVCCTDGRCCPTLQRTVYKMRFCEDGEMFSKNVMIOSCCN 327
QY 357 CPHNEASFRLY--SLFNDI 374
Db 328 CPDNDIFESLYYRKMYGDM 347

RESULT 10
US-08-468-847B-15
; Sequence 15, Application US/08468847B
; Patent No. 5780263
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
; TITLE OF INVENTION: Human CCN-like Growth Factor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GIFFILIAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468, 847B
; FILING DATE: 6 June 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-08-468-847B-15

Query Match 45.3%; Score 953.5; DB 1; Length 348;
Best Local Similarity 46.6%; Pred. No. 6.3e-73;
Matches 177; Conservative 61; Mismatches 103; Indels 39; Gaps 8;
QY 1 MSSSTRTLAVALTLHL-TRLAL-STCPAACHPLE-APKCAPGVGLVDDGCCGCCVCA 57
Db 1 MASVAGPISLAVLALCTRPATGDCSAQCQCAEAAPHCAGVSLVLDGGCCRCVCA 60
QY 58 KOLNEDOSKTOPDCHTNGLECNFGASSTALKGICRQASEGRPEYNSRITONGESPONC 117
Db 61 KOLGEICTERPDPCPHKGLFCDFGSPANKRIGVCTAK-DCAPCVFGSGSVTRSGFSOSC 119

QY 118 KHCCTCIDGAVGCIPLCPQELSLPNIGCNPRLVYVSGCCCEMWCDSDSIKSLDDDD 177
Db 120 KYCCTCIDGAVGCIPLCSMVRILPSPDCPPRRVXKLPBCKCKEMWCDEP----- 168
QY 178 LGLDASEVELTRNNELIAGKSSILKRL-PVGTETBRVLFNPLHAHGQKCIYOTMSQ 236
Db 169 -----KDRTAAGPALAAVRLDTEGPDPTM-----RANCLVOTEMSA 207
QY 237 CSKSCGTGISTRTVNDPNDPCEKRLVETRICVPRGQGVYSSILKKGKCSKTKSPPEVRF 296
Db 208 CSKTCGAGISTRTVNDPNDPCEKRLVETRICVPRGQGVYSSILKKGKCSKTKSPPEVRF 267
QY 297 TYAGCSSLVKKRYPRKYGSCVDGRCCPTLQRTVYKMRFCEDGEMFSKNVMIOSCCN 356
Db 268 ELSCGTSVKTYRAKFCGVCCTDGRCCPTLQRTVYKMRFCEDGEMFSKNVMIOSCCN 327
QY 357 CPHNEASFRLY--SLFNDI 374
Db 328 CPDNDIFESLYYRKMYGDM 347

RESULT 11
US-09-142-569-6
; Sequence 6, Application US/09142569
; Patent No. 6413735
; GENERAL INFORMATION:
; APPLICANT: Lau, Lester F.
; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Muiray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; City: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/142, 569
; FILING DATE: 02-Apr-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28758/33766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 348 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "Psp12 amino acid sequence"
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-142-569-6

Query Match 45.3%; Score 953.5; DB 4; Length 348;
Best Local Similarity 46.6%; Pred. No. 6.3e-73;
Matches 177; Conservative 61; Mismatches 103; Indels 39; Gaps 8;
QY 1 MSSSTRTLAVALTLHL-TRLAL-STCPAACHPLE-APKCAPGVGLVDDGCCGCCVCA 57
Db 1 MASVAGPISLAVLALCTRPATGDCSAQCQCAEAAPHCAGVSLVLDGGCCRCVCA 60

```

QY 58 KOLNECSTKOPDHHKGLCEGFRGASSTALIGICAOSEGRPCYSNRTRONSGSPNC 117
Db 61 KQJGELCTEDPDCHPKGLFCDFGSPANKRIGVCTAK -DGAFCVGGSVTRSGSEFSSC 119
QY 118 KHCCTCIDAVGCIPLCPQELSLPMIAGCPNRLKVSQCCSEFWVCDEDSIKDSLDDODD 177
Db 120 KYGCTCLDGAVCVPLCSMDVRLSPDPCFPFRRYKLPKCKCKEAVCDEP----- 168
QY 178 LIGLDASEVELTRNNELLAIGKSSKLRL -PVFGTEPRVYLEFNPLAHGQKCIYOTTSWSQ 236
Db 169 -----KDRATVGPALAAVRLIEDTFGDPDTM-----RANCLVQTTESMA 207
QY 237 CSKSCSTGISTRTVNDNPEBCRLVKTRTRICEVRPCGQPVYSSLKKGKCKSKTKSPSEVRF 296
Db 208 CSKTCMGISTRTVNDNTNFCRLCKCMVRPEADLEENIKKGKCKIRTPKAKEVKF 267
QY 297 TYVGGSSVKKYRPKVCGSCVDGRCTPLOTFTVKMRKPECEGSEFKSNVMIMOSCKNYN 356
Db 268 ELISGCTSVATRYAKRFEGVCTDGRCTPHRTTLLPVEKCPGELTKRNMMEIKTCACHYN 327
QY 357 CPAPHNEASRLY -SLFNDI 374
Db 328 CPGDNDIFESLYRKMYGDM 347

```

RESULT 12
US-08-167-628-2

```

GENERAL INFORMATION:
APPLICANT: Grotenodst, Gary R.
APPLICANT: Bradham Jr., Douglas M.,
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Spensley Horn Jubas & Lubitz
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/167,628
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/752,427
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr. Ph.D., John W.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: PD-1294
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-455-5100
TELEFAX: 619-455-5110
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 349 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
OS-08-167-628-2

```

Query Match	45.2%	Score	950;	DB 1:	Length	349;			
Best Local Similarity	45.7%	Pred. NC	1.2e-72;						
Matches	176;	Conservative	64;	Mismatches	97;	Indels	48;	Gaps	9;

```

QY 1 MSSSTRJIAVAATLTLHLRSLALSLCPAA-----CHCPLE-APKCAPGVGLVDDCGC 52
Db 1 MTAASGAPVRAVAVYL-----LALCSRAVGVQWCSGPCRCPDEPARRCAGAGSLVLDCCG 56
QY 53 CKVCAKOLNEDCSKTQPCDHTKGLEBCNFGASSTALKGICRAOSGRPCPEYNSRIYONGES 112
Db 57 CRYCAKOLIELCTERPDPCDHPHGLFCDFGSPANRKIGVCTAK-IGDAPCIFGTVYRSGES 115
QY 113 FQPNCKHOCTCIDGANGCIPLOPDELSFJNLCGRPRILVYKSGGCCPEWWDDEDSINDSL 172
Db 116 FQSSCKYQCTCIDGAAGCMLCSMVRLPSDPCPPRRVKILPGKCEEWVWDEK----- 170
QY 173 DDDDDLLGDASEVELTRNNELTALIGKSSSLKRL-IVRGTEPRVLEFNJLHAOGKCIYOT 231
Db 171 -DQ-----TVGPRALAAIRLEDFTGPRPTM-----RANCLVOT 203
QY 232 TSMQSCSKSGTGISTRYTNDNPECRLYKETRICEVRBCGQPVYSSLKGKCSKYTKSP 291
Db 204 TEMSACSKTCGMQISTRYTNDNASCRLKQSRBLCAVRBCEDADLEENIKGKKCIRTPKIS 263
QY 292 EPRPRTFYAGCSSYKATTRPKYICSSCYVDGRCTPLQIRYKAMFRCEDEGEMFSKNMMIQC 35
Db 264 KPIKFEISGCTSMKATYRAKFCGVCDGRCTPHRTTTLVBEFKCPDGEVMMKNMMFKTC 323
QY 352 KCNYNCPHPNEASFRLY--SLFNDI 374
Db 324 ACHYNCPCGDNDFESLITYRKMTGDM 348

```

RESULT 13
US-08-386-680-2

```

; GENERAL INFORMATION:
; APPLICANT: Grotendorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lublitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/386,680
; FILING DATE: 10-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,628
; FILING DATE:
; APPLICATION NUMBER: US/07/752,427
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-455-5100
; TELEFAX: 619-455-5110
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 349 amino acids
; type: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-386-680-2

```

Query Match 45.2%; Score 950; DB 1; Length 349;
Best Local Similarity 45.7%; Pred. No. 1.2e-72;
Matches 176; Conservative 64; Mismatches 97; Indels 48; Gaps 9;

```
QY 1 MSSSTFTLAVATLHLTRALSTCPAA-----CHCPLE-APKCAPGVGLVRCGCC 52
D 1 MTAASMPVAVAVL-----LALCSRPAVQNGSGRCRCDDEAPRCAGVSLVLDGCC 56
QY 53 CKVCAKOLNEDCKTOPCDHTKGLCECNFASSTALKICRAOSEGRCEYNSRIYNGES 112
D 57 CRYCAKOLGELCTERPCDHPKHGLFCDGSPANKKIGVCTAK-DGAPCITGGTVYRSGES 115
QY 113 FOPNCKHOCCTIDGAVGICPLCQDELSTPLNGCPNRLVYSGQCCSEWYCDSDSIKSL 172
D 116 FOSSCKYQCTCLDGAVGCMLCSMDVRLSPDCFPFRVRLPKGCCSEWYCDSEPK----- 170
QY 173 DDQDDLGLDASEVELTRNNELIAGKSSILKRL-PVFGTEPRVLFNPLAHGOKCIYOT 231
D 171 -DQ-----TVGPAALAAVRLDFTFGDPTMI-----RANCLVOT 203
QY 232 TSWSCSKSCGTGISTRTVNDNPECRVKEKTRICEVAPCGOPYSSLKKGKSKTKKSP 291
D 204 TENSACSKTCGMGISTRTVNDNASCRLKESRLCMVRPCADLEENIKKKGKCIPTPKIS 263
QY 292 EPRVFTYAGSSVKKRPRKYGSCVDGRCTPLQTRTVKMRFCDEGEMFSKNVMIQSC 351
D 264 KPIKFEISGCTSKMTYRAKFCGVCTDGRCTPHRTTLVPEFKCPDGEVAKKMMFKITC 323
QY 352 KCYNCPHPNEASFRLY--SLFNDI 374
D 324 ACHYNCBGNDIDFESLYRKMGDM 348
```

RESULT 14

US-08-459-717-2
; Sequence 2, Application US/08459717
; Patent No. 5770209
; GENERAL INFORMATION:
; APPLICANT: Grotenhorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,717
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/752,427
; FILING DATE: 30-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294
; TELEPHONE: 619-455-5100
; TELEFAX: 619-455-5110
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 349 amino acids
; TYPE: amino acid

TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-459-717-2

Query Match 45.2%; Score 950; DB 1; Length 349;
Best Local Similarity 45.7%; Pred. No. 1.2e-72;
Matches 176; Conservative 64; Mismatches 97; Indels 48; Gaps 9;

```
QY 1 MSSSTFTLAVATLHLTRALSTCPAA-----CHCPLE-APKCAPGVGLVRCGCC 52
D 1 MTAASMPVAVAVL-----LALCSRPAVQNGSGRCRCDDEAPRCAGVSLVLDGCC 56
QY 53 CKVCAKOLNEDCKTOPCDHTKGLCECNFASSTALKICRAOSEGRCEYNSRIYNGES 112
D 57 CRYCAKOLGELCTERPCDHPKHGLFCDGSPANKKIGVCTAK-DGAPCITGGTVYRSGES 115
QY 113 FOPNCKHOCCTIDGAVGICPLCQDELSTPLNGCPNRLVYSGQCCSEWYCDSDSIKSL 172
D 116 FOSSCKYQCTCLDGAVGCMLCSMDVRLSPDCFPFRVRLPKGCCSEWYCDSEPK----- 170
QY 173 DDQDDLGLDASEVELTRNNELIAGKSSILKRL-PVFGTEPRVLFNPLAHGOKCIYOT 231
D 171 -DQ-----TVGPAALAAVRLDFTFGDPTMI-----RANCLVOT 203
QY 232 TSWSCSKSCGTGISTRTVNDNPECRVKEKTRICEVAPCGOPYSSLKKGKSKTKKSP 291
D 204 TENSACSKTCGMGISTRTVNDNASCRLKESRLCMVRPCADLEENIKKKGKCIPTPKIS 263
QY 292 EPRVFTYAGSSVKKRPRKYGSCVDGRCTPLQTRTVKMRFCDEGEMFSKNVMIQSC 351
D 264 KPIKFEISGCTSKMTYRAKFCGVCTDGRCTPHRTTLVPEFKCPDGEVAKKMMFKITC 323
QY 352 KCYNCPHPNEASFRLY--SLFNDI 374
D 324 ACHYNCBGNDIDFESLYRKMGDM 348
```

RESULT 15

US-08-712-302-2
; Sequence 2, Application US/08712302
; Patent No. 5783187
; GENERAL INFORMATION:
; APPLICANT: Grotenhorst, Gary R.
; APPLICANT: Bradham Jr., Douglas M.,
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,302
; FILING DATE: 11-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,680
; FILING DATE: 10-FEB-1995
; APPLICATION NUMBER: US/08/167,628
; FILING DATE:
; APPLICATION NUMBER: US/07/752,427
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr. Ph.D., John W.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-1294

TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-455-5100
TELEFAX: 619-455-5110
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 349 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-712-302-2

Query Match 45.2%; Score 950; DB 1; Length 349;
Best Local Similarity 45.7%; Pred. No. 1.2e-72;
Matches 176; Conservative 64; Mismatches 97; Indels 48; Gaps 9;

QY 1 MSSSTFRITLAVAVTLHLRLALSTCPAA-----CHCPLE-APKCAPGVGLVRDCCG 52
Db 1 MTAASKGPYRAFAVVL---LALCSRPAVGQNCSPRCRCPDEPAPRCAPAGVSLVLDGCGC 56
QY 53 CRYCAKQLNEDCSKTQPCDHTFKGLKCNFGASSTALKGICRAOSEGRPCYNSRIYONGES 112
Db 57 CRYCAKQLNEDCSKTQPCDHTFKGLKCNFGASSTALKGICRAOSEGRPCYNSRIYONGES 115
QY 113 FQPNCKHQCCTIDGAVGCIPLCPQELSLPNLGCNPRLVYVSGQCCCEWVCDSDSIKDSL 172
Db 116 FQSCCKYQCTCLDGAAGCMPLCSMDVRLSPDCPFPRVYKLPQKCCCEWVCDSDSIKDSL 170
QY 173 DDQDLILGIDASEVELTRNNELIATGKSSLKRL-PVFTSEPRVLFNPLHAHQKCIYOT 231
Db 171 -DQ-----TVGSPALAAAYRLIEDTFGPDPTMI-----RANCLYOT 203
QY 232 TSMSCSKSGGTGISTRTVNDNPECHLYKTRICEVRCGOPYSSLSLKKKCSKTKSP 291
Db 204 TEMSASCSTCGMGISTRTVNDNASCRLKQSRCLCMVRCPEADLEENIKKGCIRTPKIS 263
QY 292 EPVRTFYAGCSSVKKYRPKYCGSCVDGRCCTPLQRTYKMRFCEDGEMFSKNVMIIQSC 351
Db 264 KPIKFELSGCTSMKTYRAKFCGVCYTDGRCCTPHRTTLPLVEFRCPDGEVKKNMFIKTC 323
QY 352 KCNYNCPHPNEASFRLY--SLFNDI 374
Db 324 ACHYNCPGDNDFESLYRKMAYDM 348

Search completed: August 5, 2003, 14:08:01
Job time : 14.9632 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:06:57 ; Search time 34.4092 Seconds

(Without alignments)
1308.079 Million cell updates/sec

Title: US-09-495-448a-2

Perfect score: 2103

Sequence: 1 MSSSTFRTLAVALTLHLTR.....PNEASFRLYSLFNDIHKFRD 379

Scoring table:

BLAST62

Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubppa/PTCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/1/pubppa/PTCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubppa/US09A_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubppa/US09B_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep:*

13: /cgn2_6/ptodata/1/pubppa/US10A_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubppa/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep:*

17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2103	100.0	379	9 US-09-853-625B-11	Sequence 11, Appl
2	2103	100.0	379	14 US-10-053-753-2	Sequence 2, Appl
3	1938	92.2	381	11 US-09-901-910-2	Sequence 2, Appl
4	1938	92.2	381	15 US-10-294-796-2	Sequence 2, Appl
5	1938	92.2	455	9 US-09-925-301-1432	Sequence 1432, Ap
6	1930	91.8	381	15 US-10-205-823-84	Sequence 84, Appl
7	1929	91.7	381	14 US-10-053-753-4	Sequence 4, Appl
8	1819.5	86.5	374	9 US-09-853-625B-12	Sequence 12, Appl
9	1819.5	86.5	375	11 US-09-901-910-7	Sequence 7, Appl
10	1646.5	78.3	375	9 US-09-853-625B-13	Sequence 13, Appl
11	960	45.6	347	15 US-10-245-977-7	Sequence 7, Appl
12	957.5	45.5	348	14 US-10-101-040-3	Sequence 3, Appl
13	953.5	45.3	348	9 US-09-853-625B-15	Sequence 15, Appl
14	953.5	45.3	348	14 US-10-053-753-6	Sequence 6, Appl
15	953.5	45.3	348	15 US-10-245-977-8	Sequence 8, Appl

16	950	45.2	349	14 US-10-101-040-4	Sequence 4, Appl
17	950	45.2	349	14 US-10-011-859-26	Sequence 26, Appl
18	950	45.2	349	14 US-10-053-753-8	Sequence 8, Appl
19	950	45.2	349	15 US-10-060-036-173	Sequence 173, Appl
20	950	45.2	349	15 US-10-171-311-46	Sequence 46, Appl
21	950	45.2	349	15 US-10-205-823-78	Sequence 78, Appl
22	950	45.2	349	15 US-10-245-977-2	Sequence 2, Appl
23	944	44.9	347	14 US-10-080-173-2	Sequence 2, Appl
24	944	44.9	347	14 US-10-101-040-2	Sequence 2, Appl
25	941.5	44.8	348	9 US-09-853-625B-14	Sequence 14, Appl
26	937	44.6	349	15 US-10-245-977-6	Sequence 6, Appl
27	919	43.7	349	15 US-10-245-977-7	Sequence 7, Appl
28	861	40.9	351	9 US-09-853-625B-16	Sequence 16, Appl
29	839.5	39.9	357	9 US-09-853-625B-17	Sequence 17, Appl
30	839.5	39.9	357	14 US-10-011-859-25	Sequence 25, Appl
31	773	36.8	345	15 US-10-112-267-3	Sequence 3, Appl
32	773	36.8	345	15 US-10-112-267-6	Sequence 6, Appl
33	773	36.8	345	15 US-10-112-267-5	Sequence 5, Appl
34	773	36.8	367	15 US-10-001-054-50	Sequence 50, Appl
35	773	36.8	367	15 US-10-112-267-4	Sequence 4, Appl
36	772	36.7	345	15 US-10-112-267-8	Sequence 8, Appl
37	772	36.7	345	15 US-10-112-267-5	Sequence 5, Appl
38	772	36.7	367	15 US-10-112-267-21	Sequence 21, Appl
39	772	36.7	367	15 US-10-112-267-7	Sequence 7, Appl
40	758.5	36.1	345	15 US-10-112-267-22	Sequence 22, Appl
41	758.5	36.1	367	15 US-10-112-267-11	Sequence 11, Appl
42	622	29.6	339	15 US-10-112-267-36	Sequence 36, Appl
43	622	29.6	354	14 US-10-011-859-2	Sequence 2, Appl
44	622	29.6	354	14 US-10-112-267-37	Sequence 37, Appl
45	621	29.5	339	15 US-10-112-267-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1

US-09-853-625B-11

Sequence 11, Application US/09853625B

Patent No. US20020049304A1

GENERAL INFORMATION:

APPLICANT: Haslings, Gregg A. and Adams, Mark D.

TITLE OF INVENTION: Human CCN-Like Growth Factor

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESSES:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILPILLIAN,

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/853, 625B

FILING DATE: 14-May-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.

REGISTRATION NUMBER: 33, 073

REFERENCE/DOCKET NUMBER: 325800-442

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 379 AMINO ACIDS

TYPE: AMINO ACID
STRANDEDNESS: <Unknown>
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-653-6258-11

Query Match 100.0%; Score 2103; DB 9; Length 379;
Best Local Similarity 100.0%; Pred. No. 3.2e-167;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSTFTLVAAYTLHLTRALSTCPAACHCPLKAPGVGLVDDGGCCCKVCAKOL 60
DB 1 MSSSTFTLVAAYTLHLTRALSTCPAACHCPLKAPGVGLVDDGGCCCKVCAKOL 60
QY 61 NEDSKTOPCDHTKGLKCNFGASTALGICRAOSEGRPEYNSRIYONGESFQPNCKHQ 120
DB 61 NEDSKTOPCDHTKGLKCNFGASTALGICRAOSEGRPEYNSRIYONGESFQPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNIGCPNPLVKVSGCCCEWVCDSDSIRKSLDDQDILLG 180
DB 121 CTCIDGAVGCIPLCPQELSLPNIGCPNPLVKVSGCCCEWVCDSDSIRKSLDDQDILLG 180
QY 181 LDASEVELTRNNELIAGKSSSLKRLPVFETEPRLVFNPLHAHQKCIYQTTSMSCSKS 240
DB 181 LDASEVELTRNNELIAGKSSSLKRLPVFETEPRLVFNPLHAHQKCIYQTTSMSCSKS 240
QY 241 CGGISTRTVNDNPECLVETRICVRCGOPYSSLKKGKCKSTKKSPEVPRTYAG 300
DB 241 CGGISTRTVNDNPECLVETRICVRCGOPYSSLKKGKCKSTKKSPEVPRTYAG 300
QY 301 CSSVKYKRPYKCGSCVDGRCTPLQRTVYKMRFCEDGEMFSKNVMYIOSCKCNYNCPHP 360
DB 301 CSSVKYKRPYKCGSCVDGRCTPLQRTVYKMRFCEDGEMFSKNVMYIOSCKCNYNCPHP 360
QY 361 NEASFRLYSLENDIHKFRD 379
DB 361 NEASFRLYSLENDIHKFRD 379

RESULT 2 US-10-053-753-2

; Sequence 2, Application US/10053753
; Publication No. US20020150986A1

GENERAL INFORMATION:

; APPLICANT: Lau, Lester F.
; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/053,753
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 28758/33766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:

NAME/KEY: misc-feature
OTHER INFORMATION: "Mouse Cyr61 amino acid sequence"
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-053-753-2

Query Match 100.0%; Score 2103; DB 14; Length 379;
Best Local Similarity 100.0%; Pred. No. 3.2e-167;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSTFTLVAAYTLHLTRALSTCPAACHCPLKAPGVGLVDDGGCCCKVCAKOL 60
DB 1 MSSSTFTLVAAYTLHLTRALSTCPAACHCPLKAPGVGLVDDGGCCCKVCAKOL 60
QY 61 NEDSKTOPCDHTKGLKCNFGASTALGICRAOSEGRPEYNSRIYONGESFQPNCKHQ 120
DB 61 NEDSKTOPCDHTKGLKCNFGASTALGICRAOSEGRPEYNSRIYONGESFQPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNIGCPNPLVKVSGCCCEWVCDSDSIRKSLDDQDILLG 180
DB 121 CTCIDGAVGCIPLCPQELSLPNIGCPNPLVKVSGCCCEWVCDSDSIRKSLDDQDILLG 180
QY 181 LDASEVELTRNNELIAGKSSSLKRLPVFETEPRLVFNPLHAHQKCIYQTTSMSCSKS 240
DB 181 LDASEVELTRNNELIAGKSSSLKRLPVFETEPRLVFNPLHAHQKCIYQTTSMSCSKS 240
QY 241 CGGISTRTVNDNPECLVETRICVRCGOPYSSLKKGKCKSTKKSPEVPRTYAG 300
DB 241 CGGISTRTVNDNPECLVETRICVRCGOPYSSLKKGKCKSTKKSPEVPRTYAG 300
QY 301 CSSVKYKRPYKCGSCVDGRCTPLQRTVYKMRFCEDGEMFSKNVMYIOSCKCNYNCPHP 360
DB 301 CSSVKYKRPYKCGSCVDGRCTPLQRTVYKMRFCEDGEMFSKNVMYIOSCKCNYNCPHP 360
QY 361 NEASFRLYSLENDIHKFRD 379
DB 361 NEASFRLYSLENDIHKFRD 379

RESULT 3 US-09-901-910-2

; Sequence 2, Application US/09901910
; Publication No. US20030012768A1

GENERAL INFORMATION:

; APPLICANT: Li, Haodong
; APPLICANT: Adams, Mark
; APPLICANT: Calenda Valerie
; TITLE OF INVENTION: Connective Tissue Growth Factor-2
; FILE REFERENCE: PFI26P2
; CURRENT APPLICATION NUMBER: US/09/901,910
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/348,815
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 08/459,101
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: PCT/US94/07736
; PRIOR FILING DATE: 1994-07-12
; PRIOR APPLICATION NUMBER: 60/217,402
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/291,642
; PRIOR FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 2
; LENGTH: 381
; TYPE: prt

; ORGANISM: homo sapiens
US-09-901-910-2

Query Match 92.2%; Score 1938; DB 11; Length 381;
Best Local Similarity 91.4%; Pred. No. 1.8e-153;
Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSSTFTLAVAVTLHLTRIALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
DB 1 MSSRIARALAVTLHLTRIALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
QY 61 NEDCSKTOPCDHTKGLGECNFGASSTALKGICRAOSEGRPCENSRITYONGESFOPNCKHQ 120
DB 61 NEDCSKTOPCDHTKGLGECNFGASSTALKGICRAOSEGRPCENSRITYONGESFOPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVYKVGQCCCEWVCDSDSIKSLDDODDL- 178
DB 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVYKVGQCCCEWVCDSDSIKSLDDODDL 180
QY 179 --LGIDASEVELTRNNELIAGKSSILKRLPVFGTEPRVLFNPLAHGOKCIYOTTSMQ 236
DB 181 KELGFDASEVELTRNNELIAGKSSILKRLPVFGTEPRVLFNPLAHGOKCIYOTTSMQ 238
QY 237 CSKSCGIGISIRVTNNDNEPCRLVETRICERPCGOPYSSSLKGGKSKTKKSPPEYRF 296
DB 239 CSKTCGTGISTRTVNDNEPCRLVETRICERPCGOPYSSSLKGGKSKTKKSPPEYRF 298
QY 297 TYACSSVKKYRPYKCGSCVDGRCTPLQRTVYKMRFCEDGEMFSKNVMMIIOCKCNYN 356
DB 299 TYACGLSVKRYRPYKCGSCVDGRCTPLQRTVYKMRFCEDGEMFSKNVMMIIOCKCNYN 358
QY 357 CPHNEASFRLYSLEFNDIHKFRD 379
DB 359 CPHANEAFPEYRLFNDIHKFRD 381

RESULT 4

US-10-294-796-2
; Sequence 2, Application US/10294796
; Publication No. US20030078391A1
; GENERAL INFORMATION:
; APPLICANT: Li, Haodong et al.
; TITLE OF INVENTION: Connective Tissue Growth Factor-2
; FILE REFERENCE: P126P1D2
; CURRENT APPLICATION NUMBER: US/10/294,796
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 09/348,815
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US 08/459,101
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: PCT/US94/07736
; PRIOR FILING DATE: 1994-07-12
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-796-2

Query Match 92.2%; Score 1938; DB 15; Length 381;
Best Local Similarity 91.4%; Pred. No. 1.8e-153;
Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSSTFTLAVAVTLHLTRIALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
DB 1 MSSRIARALAVTLHLTRIALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
QY 61 NEDCSKTOPCDHTKGLGECNFGASSTALKGICRAOSEGRPCENSRITYONGESFOPNCKHQ 120
DB 61 NEDCSKTOPCDHTKGLGECNFGASSTALKGICRAOSEGRPCENSRITYONGESFOPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVYKVGQCCCEWVCDSDSIKSLDDODDL- 178
DB 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVYKVGQCCCEWVCDSDSIKSLDDODDL 180

QY 179 --LGIDASEVELTRNNELIAGKSSILKRLPVFGTEPRVLFNPLAHGOKCIYOTTSMQ 236
DB 181 KELGFDASEVELTRNNELIAGKSSILKRLPVFGTEPRVLFNPLAHGOKCIYOTTSMQ 238
QY 237 CSKSCGIGISIRVTNNDNEPCRLVETRICERPCGOPYSSSLKGGKSKTKKSPPEYRF 296
DB 239 CSKTCGTGISTRTVNDNEPCRLVETRICERPCGOPYSSSLKGGKSKTKKSPPEYRF 298
QY 297 TYACSSVKKYRPYKCGSCVDGRCTPLQRTVYKMRFCEDGEMFSKNVMMIIOCKCNYN 356
DB 299 TYACGLSVKRYRPYKCGSCVDGRCTPLQRTVYKMRFCEDGEMFSKNVMMIIOCKCNYN 358
QY 357 CPHNEASFRLYSLEFNDIHKFRD 379
DB 359 CPHANEAFPEYRLFNDIHKFRD 381

RESULT 5

US-09-925-301-1432
; Sequence 1432, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P1106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1432
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1432

Query Match 92.2%; Score 1938; DB 9; Length 455;
Best Local Similarity 91.4%; Pred. No. 2.2e-153;
Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSSTFTLAVAVTLHLTRIALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
DB 75 MSSRIARALAVTLHLTRIALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 134
QY 61 NEDCSKTOPCDHTKGLGECNFGASSTALKGICRAOSEGRPCENSRITYONGESFOPNCKHQ 120
DB 135 NEDCSKTOPCDHTKGLGECNFGASSTALKGICRAOSEGRPCENSRITYONGESFOPNCKHQ 194
QY 121 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVYKVGQCCCEWVCDSDSIKSLDDODDL- 178
DB 195 CTCIDGAVGCIPLCPQELSLPNIIGCPNRLVYKVGQCCCEWVCDSDSIKSLDDODDL 254
QY 179 --LGIDASEVELTRNNELIAGKSSILKRLPVFGTEPRVLFNPLAHGOKCIYOTTSMQ 236
DB 255 KELGFDASEVELTRNNELIAGKSSILKRLPVFGTEPRVLFNPLAHGOKCIYOTTSMQ 312
QY 237 CSKSCGIGISIRVTNNDNEPCRLVETRICERPCGOPYSSSLKGGKSKTKKSPPEYRF 296
DB 313 CSKTCGTGISTRTVNDNEPCRLVETRICERPCGOPYSSSLKGGKSKTKKSPPEYRF 372
QY 297 TYACSSVKKYRPYKCGSCVDGRCTPLQRTVYKMRFCEDGEMFSKNVMMIIOCKCNYN 356
DB 373 TYACGLSVKRYRPYKCGSCVDGRCTPLQRTVYKMRFCEDGEMFSKNVMMIIOCKCNYN 432
QY 357 CPHNEASFRLYSLEFNDIHKFRD 379
DB 433 CPHANEAFPEYRLFNDIHKFRD 455

RESULT 6
US-10-205-823-84

Sequence 84, Application US/10205823
Publication No. US20030108963A1

GENERAL INFORMATION:

APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Endege, Wilson O.
APPLICANT: Gannavarapu, Manjula
APPLICANT: Gorbacheva, Bella
APPLICANT: Hoersch, Sebastian
APPLICANT: Kamatkar, Shubhang
APPLICANT: Monsey, Angela M.
APPLICANT: Ghatt, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Anderson, Dustin
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE OF INVENTION: THERAPY OF PROSTATE CANCER
FILE REFERENCE: MRI-044
CURRENT APPLICATION NUMBER: US/10/205, 823
CURRENT FILING DATE: 2002-07-25
PRIOR APPLICATION NUMBER: 60/307, 982
PRIOR FILING DATE: 2001-07-25
PRIOR APPLICATION NUMBER: 60/314, 356
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/325, 020
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: 60/341, 746
PRIOR FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: 60/362, 158
PRIOR FILING DATE: 2002-03-05
NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 84
LENGTH: 381
TYPE: PRT
ORGANISM: Homo sapiens
US-10-205-823-84

Query Match

Best Local Similarity 91.8%; Score 1930; DB 15; Length 381;
Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSSSTFTLAVAVTLHLRLALSTCPACCHPCLAPKCAPGVGLVDRDGGCCCKVCAKOL 60
DB 1 MSSRIKALALVYTLHLRLALSTCPACCHPCLAPKCAPGVGLVDRDGGCCCKVCAKOL 60
QY 61 NEDCSKTOPCDHTKGLCECNFGASSTALKGICRAQSGRCPEYNSRIYONGESFOPNCKHQ 120
DB 61 NEDCSKTOPCDHTKGLCECNFGASSTALKGICRAQSGRCPEYNSRIYONGESFOPNCKHQ 120
QY 121 CTCIDAVAGCIPCLPELSPNLCGPNRLVYKSGGCCCEWVDEDSINDSDDDDL-- 178
DB 121 CTCIDAVAGCIPCLPELSPNLCGPNRLVYKSGGCCCEWVDEDSINDSDDDDLG 180
QY 179 --LGLDASEVELLRNNELLAIGKSSSLKRLPVGTEPRVLFNPLHAHGOKCIYQTTSMQ 236
DB 181 KELGFDASEVELLRNNELLAIGKSSSLKRLPVGTEPRVLFNPLHAHGOKCIYQTTSMQ 238
QY 237 CSKSGCTGISTRTVNDNPECRVLKTRICEVPRCGQPVYSSLLKGGKSKTKKSPBPVRF 296
DB 239 CSKSGCTGISTRTVNDNPECRVLKTRICEVPRCGQPVYSSLLKGGKSKTKKSPBPVRF 298
QY 297 TYAGCSSVKKYRKRYGSCVDGRCCPLOTRTYKMFRCEDGEMFSKNVMIMOSCKCNYN 356
DB 299 TYAGCSSVKKYRKRYGSCVDGRCCPLOTRTYKMFRCEDGEMFSKNVMIMOSCKCNYN 358
QY 357 CPHPNEASFRLYSLFNIDHKFRD 379
DB 359 CPHPNEASFRLYSLFNIDHKFRD 381

RESULT 7
US-10-053-753-4

Sequence 4, Application US/10053753
Publication No. US20020150986A1

GENERAL INFORMATION:

APPLICANT: Lau, Lester F.
TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/053, 753
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28758/33766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 381 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: "human Cyr61 amino acid sequence"
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-053-753-4

Query Match

Best Local Similarity 91.7%; Score 1929; DB 14; Length 381;
Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSSSTFTLAVAVTLHLRLALSTCPACCHPCLAPKCAPGVGLVDRDGGCCCKVCAKOL 60
DB 1 MSSRIKALALVYTLHLRLALSTCPACCHPCLAPKCAPGVGLVDRDGGCCCKVCAKOL 60
QY 61 NEDCSKTOPCDHTKGLCECNFGASSTALKGICRAQSGRCPEYNSRIYONGESFOPNCKHQ 120
DB 61 NEDCSKTOPCDHTKGLCECNFGASSTALKGICRAQSGRCPEYNSRIYONGESFOPNCKHQ 120
QY 121 CTCIDAVAGCIPCLPELSPNLCGPNRLVYKSGGCCCEWVDEDSINDSDDDDL-- 178
DB 121 CTCIDAVAGCIPCLPELSPNLCGPNRLVYKSGGCCCEWVDEDSINDSDDDDLG 180
QY 179 --LGLDASEVELLRNNELLAIGKSSSLKRLPVGTEPRVLFNPLHAHGOKCIYQTTSMQ 236
DB 181 KELGFDASEVELLRNNELLAIGKSSSLKRLPVGTEPRVLFNPLHAHGOKCIYQTTSMQ 238
QY 237 CSKSGCTGISTRTVNDNPECRVLKTRICEVPRCGQPVYSSLLKGGKSKTKKSPBPVRF 296
DB 239 CSKSGCTGISTRTVNDNPECRVLKTRICEVPRCGQPVYSSLLKGGKSKTKKSPBPVRF 298
QY 297 TYAGCSSVKKYRKRYGSCVDGRCCPLOTRTYKMFRCEDGEMFSKNVMIMOSCKCNYN 356
DB 299 TYAGCSSVKKYRKRYGSCVDGRCCPLOTRTYKMFRCEDGEMFSKNVMIMOSCKCNYN 358
QY 357 CPHPNEASFRLYSLFNIDHKFRD 379


```

: PRIOR APPLICATION NUMBER: 09/292, 036
: PRIOR FILING DATE: 1999-04-14
: PRIOR APPLICATION NUMBER: US 09/292, 036
: PRIOR FILING DATE: 1999-04-14
: PRIOR APPLICATION NUMBER: US 09/187, 478
: PRIOR FILING DATE: 1998-11-06
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO: 3
: LENGTH: 348
: TYPE: PRT
: ORGANISM: Mouse
US-10-101-040-3

```

Query Match	45.5%	Score 957.5	DB 14	length 348
Best Local Similarity	46.8%	Pred. No. 7e-72		
Matches 178; Conservative	60;	Mismatches 103;	Indels 39;	Gaps 8;

[illegible]

RESULT 13
US-09-853-625B-15
Sequence 15, Application US/09853625B
Patent No. US20020049304A1
GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factors
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/853,625B
FILING DATE: 14-May-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/053,587
 FILING DATE: <unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: MULLINS, J G
 REGISTRATION NUMBER: 33,073
 REFERENCE/DOCKET NUMBER: 325800-442
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 348 AMINO ACIDS
 TYPE: AMINO ACID
 STRANDEDNESS: <Unknown>
 TOPOLOGY: LINEAR
 MOLECULE TYPE: PROTEIN
 SEQUENCE DESCRIPTION: SEQ ID NO: 15:
 MS-09-853-625b-15

Query Match	45.3%	Score 953.5	DB 9	Length 348
Best Local Similarity	46.6%	Pred. No. 1.5e-71		
Matches 177	Conservative 61	Mismatches 103	Indels 39	Gaps 8

QY	1	MSSSTFRRLTAAVTLILH-TREAL-STCPACHQPLE-APKCAFGVGLVRGGCCCKVCA	57
Db	1	MAASVAGPISLIALVLALCTTPATGODCSAQOQCAEAEPAPGAVSLVLDGGCCCRKA	60
QY	58	KQLNEDCSSTQPCDHTKGLCEMFGASSTALKGICRAQSEGRPCENSRHXYONGESFQPN	117
Db	61	KQGLCTCTERDPCDHPKGLGFQDGPAPNARKIGVCTAK-DGAPCVAGSGSVASNGSPSSC	119
QY	118	KKQCTCTIDAGVCIPLCPQELSLPMLGCPNRLKAVSGQCCBEWVCDEDSLKSLDDDD	177
Db	120	KYQCTCIDAVGCVPLCSMDYRLPDPDPPFRARVLLPCKCKEAWCDEP-----	168
QY	178	LGLDASEVELRNNEMLAIGKSGSLKLRL-PVFGEPRLVFNPLHAGQKCIYQVTSMSQ	236
Db	169	-----KDRATVGPALAAAYRLDFTFGDPITM-----RANCLVOTTEMSA	207
QY	237	CSKSCGTGISTRYTNDNPECRLYKETRICEVAPGQPYSSLKKGKCSKTKSKPEVRF	296
Db	208	CSKTCGTGISTRYTNDNFRCRLKLEKSSRLCQVAPCADLEINKKGKCIIRPKIAKAYKE	267
QY	297	TYAGCSYVKKRPKICGSCVDRGCTPLQIOTRYVMKRFCEDEGMFSKNVMMIOSCKCNY	356
Db	268	ELSGCTSVTYAKRFQGVCTDGRCTCPHRTITLLPYEFKCPDGEITMKNMMEIKTCAACHY	327
QY	357	CPHPNEASFRLY-SLFPNDI 374	
Db	328	CPEDNDIEESLYKKKMGDM 347	

RESULT 14
 US-10-053-753-6
 Sequence 6, Application US/10053753
 Publication No. US20020150986v1
 GENERAL INFORMATION:
 APPLICANT: Lau, Lester F.
 TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 6300 Sears Tower, 233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/053,753
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28758/33766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: "Faspl2 amino acid sequence"
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-053-753-6

Query Match 45.3%; Score 953.5; DB 14; Length 348;
Best Local Similarity 46.6%; Pred. No. 1.5e-71;
Matches 177; Conservative 61; Mismatches 103; Indels 39; Gaps 8;

QY 1 MSSSTFTLAVATLHL-TRLAL-STCPAACHPLE-APKCAAGVGLVDDGCCCKVCA 57
1 MLASVAGPISLALVLAICTRPATGDCSACCCAAEAAPHCPAGVSLVLDGCGCCRYCA 60
QY 58 KOLNEDCKTOPCDHTKLECNFGASSTALKGICRAOSEGRPCPEYNSRIYONGESFPNC 117
61 KOLGELCTERDPCDPHKGFLCDFGSPANKRIGVCTAK-DGAPCVFGGSVRSGESFSSC 119
QY 118 KHOCTCIDGAVGCIPLCPQELSLPNLGCNPRLVYKSGCCCEWVCDSDSIRSLDDDD 177
120 KYOCTCIDGAVGCVPLCSMDVRLPSPDPPRRVKLPKCKCKEWCDEP----- 168
QY 178 LIGIDASEVELTNNELIATGKSSSLKRL-PVFETEPRLVFNPLHAHGQKCIYOTTSMSQ 236
169 -----KDRTAAGPALAAVRLDTEGPDPPTM-----RANCLVQTTWMSA 207
QY 237 CSKSCGTGISTRYTNDNPECRLYKETRICEVRCPOGQPVYSSLKKGKCKSTKSPPEVRP 296
208 CSKTCGKGISTRYTNDNTPCRLKQSLCVRPOEADLEINIKKGCIRTPIAKPVKF 267
QY 297 TYAGCSSVKYRPRKYGSCVDGRCTPLQTRTYKMRFRCDGEMFNSKNVMIOCKCNYN 356
268 ELGCGTSVKTYRAKFCGCTDGRCCPHRTTLLPVEFKCPDGEIMKKNMFIKTCACHYN 327
QY 357 CPHNEASFRLY--SLFNDI 374
DB 328 CPGDNDIFESLYTKMTGDM 347

RESULT 15

US-10-245-977-8
Sequence 8, Application US/10245977
Publication No. US20030113816A1
GENERAL INFORMATION:
APPLICANT: Weitz, Stephen L.
APPLICANT: Usinger, William R.
TITLE OF INVENTION: METHODS OF ASSAYING CONNECTIVE TISSUE GROWTH FACTOR
FILE REFERENCE: FP0812 US
CURRENT APPLICATION NUMBER: US/10/245,977
CURRENT FILING DATE: 2002-09-18
PRIOR APPLICATION NUMBER: US 60/323,305
PRIOR FILING DATE: 2001-09-18
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8

LENGTH: 348
TYPE: PRT
ORGANISM: Mus musculus
US-10-245-977-8

Query Match 45.3%; Score 953.5; DB 15; Length 348;
Best Local Similarity 46.6%; Pred. No. 1.5e-71;
Matches 177; Conservative 61; Mismatches 103; Indels 39; Gaps 8;

QY 1 MSSSTFTLAVATLHL-TRLAL-STCPAACHPLE-APKCAAGVGLVDDGCCCKVCA 57
1 MLASVAGPISLALVLAICTRPATGDCSACCCAAEAAPHCPAGVSLVLDGCGCCRYCA 60
QY 58 KOLNEDCKTOPCDHTKLECNFGASSTALKGICRAOSEGRPCPEYNSRIYONGESFPNC 117
61 KOLGELCTERDPCDPHKGFLCDFGSPANKRIGVCTAK-DGAPCVFGGSVRSGESFSSC 119
QY 118 KHOCTCIDGAVGCIPLCPQELSLPNLGCNPRLVYKSGCCCEWVCDSDSIRSLDDDD 177
120 KYOCTCIDGAVGCVPLCSMDVRLPSPDPPRRVKLPKCKCKEWCDEP----- 168
QY 178 LIGIDASEVELTNNELIATGKSSSLKRL-PVFETEPRLVFNPLHAHGQKCIYOTTSMSQ 236
169 -----KDRTAAGPALAAVRLDTEGPDPPTM-----RANCLVQTTWMSA 207
QY 237 CSKSCGTGISTRYTNDNPECRLYKETRICEVRCPOGQPVYSSLKKGKCKSTKSPPEVRP 296
208 CSKTCGKGISTRYTNDNTPCRLKQSLCVRPOEADLEINIKKGCIRTPIAKPVKF 267
QY 297 TYAGCSSVKYRPRKYGSCVDGRCTPLQTRTYKMRFRCDGEMFNSKNVMIOCKCNYN 356
268 ELGCGTSVKTYRAKFCGCTDGRCCPHRTTLLPVEFKCPDGEIMKKNMFIKTCACHYN 327
QY 357 CPHNEASFRLY--SLFNDI 374
DB 328 CPGDNDIFESLYTKMTGDM 347

Search completed: August 5, 2003, 14:18:50
Job time : 36.4092 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:03:01 ; Search time 15.4592 Seconds
(without alignments)
2357.684 Million cell updates/sec

Title: US-09-495-448A-2

Sequence: 1 MSSSTFRTLAVALTLHLTR.....PNEASRLYSLEFNDIHKFRD 379

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2103	100.0	379	2	A35669	gene CYR61 protein
2	1646.5	78.3	375	2	A41428	CEF-10 protein pre
3	957.5	45.5	348	2	A40578	beta IG-M2 protein
4	950	45.2	349	2	A40551	connective tissue
5	861	40.9	351	2	S20078	NOV protein - chic
6	839.5	39.9	357	2	I38069	gene NOV protein
7	171	8.1	1111	2	T26972	hypothetical prote
8	166	7.9	1034	2	JC5598	mucin - rat
9	157.5	7.5	13288	2	T03099	MEGF6 protein - ra
10	156.5	7.4	1574	2	T13954	mucin, submaxillar
11	156.5	7.4	1620	2	T27283	hypothetical prote
12	156	7.4	1025	2	T42636	secreted leucine-r
13	152.5	7.3	1042	2	A57534	mucin 5Ac (clone L
14	152	7.2	1056	2	A53767	mucin MUC5B, trach
15	151.5	7.2	1700	2	S08167	Balblani ring 3 pr
16	146	6.9	837	2	A42112	mucin-like peptide
17	145	6.9	601	2	T22025	hypothetical prote
18	145	6.9	601	2	D89711	protein F40E10.4 l
19	145	6.9	1531	2	T42218	slit-1 protein hom
20	144.5	6.9	1170	2	A53612	laminin B1k chain
21	143	6.8	1101	2	T09059	notch4 - mouse
22	142.5	6.8	1101	2	T16840	hypothetical prote
23	141	6.7	251	2	A55035	cysteine-rich prot
24	141	6.7	5376	2	T42215	zonadhesin - mouse
25	140	6.7	3106	1	S53868	laminin alpha-2 ch
26	139.5	6.6	1178	1	A39604	thrombospondin pre
27	139	6.6	473	2	A56175	adhesive plaque pr
28	138.5	6.6	1847	2	T18308	probable vitelloloe
29	136.5	6.5	3020	2	A43932	mucin 2 precursor,

30	135.5	6.4	305	2	I48601	insulin-like growt
31	135	6.4	1523	2	T13953	MEGF5 protein - ra
32	135	6.4	3672	2	T23433	hypothetical prote
33	135	6.4	3704	2	T37316	probable laminin a
34	134	6.4	2219	2	T27684	hypothetical prote
35	133.5	6.3	305	2	JN0508	insulin-like growt
36	133.5	6.3	424	2	S11676	spore coat protein
37	133.5	6.3	2555	2	A40043	notch protein homo
38	133.5	6.3	3002	2	A47221	fibritillin 1 precu
39	133	6.3	1220	2	A56136	jagged protein pre
40	133	6.3	2531	2	S18188	notch protein homo
41	132.5	6.3	1106	2	T18649	hypothetical prote
42	132.5	6.3	1106	2	T13938	gene shuttle craft
43	132.5	6.3	1106	2	T44598	hypothetical prote
44	132	6.3	1292	2	T09229	galactose binding
45	131.5	6.3	810	2	T10756	Nel-homology protei

ALIGNMENTS

RESULT 1

A35669

gene CYR61 protein precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 28-Sep-1990 #sequence_revision: 18-Nov-1992 #text_change: 05-Nov-1999

C:Accession: A35669; I48319; S16446

R:O'Brien, T.P.; Yang, G.P.; Sanders, L.; Lau, L.F.

Mol. Cell. Biol. 10, 3569-3577, 1990

A:Title: Expression of cyr61, a growth factor-inducible immediate-early gene.

A:Reference number: A35669; MUID:90287146; PMID:2355916

A:Accession: A35669

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-379 <OAB>

A:Cross-references: GB:M32490; NID:9192909; PIDN:AAA37512.1; PID:9309206

A>Note: the authors translated the codon GAT for residue 337 as Gln

A:Note: the authors translated the codon GAT for residue 337 as Gln

A:Gene: CYR61

A:Inserts: 21/3; 93/1; 208/1; 279/3

C:Superfamily: von Willebrand factor type C repeat homology

F:99-166/Domain: von Willebrand factor type C repeat homology <WVC>

Query Match 100.0%; Score 2103; DB 2; Length 379;

Best Local Similarity 100.0%; Pred. No. 1.5e-141;

Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSSSTFRTLAVALTLHLTRALSTCPAACHCPLEAFRCAPGVGLVDRDGCCKVCAKOL	60
DB	1	MSSSTFRTLAVALTLHLTRALSTCPAACHCPLEAFRCAPGVGLVDRDGCCKVCAKOL	60
QY	61	NEDCSKTPQCDHRTKGLCNFGASTAIGICRAOSEBPCEYNSRIYONGESFQPNCKHQ	120
DB	61	NEDCSKTPQCDHRTKGLCNFGASTAIGICRAOSEBPCEYNSRIYONGESFQPNCKHQ	120
QY	121	CTCTDGAAGCICPLCPQELSLPNLGCNPRLVYKSGGCCCEMYVCDSDIKSLDDQDPLG	180
DB	121	CTCTDGAAGCICPLCPQELSLPNLGCNPRLVYKSGGCCCEMYVCDSDIKSLDDQDPLG	180
QY	181	LDASEVELTRNNELIAGKSSSLKRLPVFGTEPRVLPNPLAHGQKCIYOTTSWQSKS	240
DB	181	LDASEVELTRNNELIAGKSSSLKRLPVFGTEPRVLPNPLAHGQKCIYOTTSWQSKS	240

QY 241 CGTGISTRTVNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKCSKTKKSPPEVFTYAG 300
 |||||
 Db 241 CGTGISTRTVNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKCSKTKKSPPEVFTYAG 300
 QY 301 CSSVKKYRRKYGCSYDGRCTPLQTRVTKMRFRCDGEMFSKNVMIOCKKNYCNCPHP 360
 |||||
 Db 301 CSSVKKYRRKYGCSYDGRCTPLQTRVTKMRFRCDGEMFSKNVMIOCKKNYCNCPHP 360
 QY 361 NEASFRLYSLENDIHKFRD 379
 |||||
 Db 361 NEASFRLYSLENDIHKFRD 379

RESULT 2

CEF-10 protein precursor - chicken
 A:Accession: A41428
 C:Species: Gallus gallus (chicken)
 C:Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 21-Jul-2000
 C:Accession: A41428

R:Simmons, D.L.; Levy, D.B.; Yannoni, Y.; Erikson, R.L.
 Proc. Natl. Acad. Sci. U.S.A. 86, 1178-1182, 1989
 A:Title: Identification of a phorbol ester-repressible v-src-inducible gene.
 A:Reference number: A41428; MID:89145206; PMID:2537491

A:Accession: A41428
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-375 <SIM>
 A:Cross-references: GB:J04496; NID:9211435; PIDN:AAA8661.1; PID:9211436

Query Match 78.3%; Score 1646.5; DB 2; Length 375;
 Best Local Similarity 80.1%; Pred. No. 2.9e-109;
 Matches 302; Conservative 22; Mismatches 36; Indels 17; Gaps 7;

QY 10 AAVATLHTLTRLAL-STCPAACHPLEAPKCAPGVGLVVDGCGCCVCAKOLNEDCSKRO 68
 ||:|:|||||
 Db 9 ALAAALLCARLALSGPCPAVCCPAAPQACAPAGVGLVVDGCGCCVCAKOLNEDCSKRO 68
 QY 69 PCDHRTGLCECNFGASSTALKGICRAQSEGRPCENSRIVONGESFOQPNCKHOCTCIDGAV 128
 |||||
 Db 69 PCDHRTGLCECNFGASSTALKGICRAQSEGRPCENSRIVONGESFOQPNCKHOCTCIDGAV 128
 QY 129 GCTPLCPQELSLPNCPCPNRLVYKVSGCCCEWVCEDESIKOSLDQDDL---LGLDAS 184
 |||||
 Db 129 GCTPLCPQELSLPNCPCPNRLVYKVSGCCCEWVCEDESIKOSLDQDDL---LGLDAS 184
 QY 185 EVELTNNNELIAGKSSSLKRLPVFETEP--RYLNPFLAHGOKICVOTTSNOCSSKSG 242
 |||||
 Db 187 EVELTNNNELIAGKSSSLKRLPVFETEP--RYLNPFLAHGOKICVOTTSNOCSSKSG 242
 QY 243 TGISTRVTNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKCSKTKKSPPEVFTYAGS 302
 |||||
 Db 240 TGISTRVTNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKCSKTKKSPPEVFTYAGS 299
 QY 303 SVKKYRRKYGCSYDGRCTPLQTRVTKMRFRCDGEMFSKNVMIOCKKNYCNCPHP 362
 |||||
 Db 300 SVKKYRRKYGCSYDGRCTPLQTRVTKMRFRCDGEMFSKNVMIOCKKNYCNCPHP 359
 QY 363 ASFRILYSLENDIHKFRD 379
 ||:|:|||||
 Db 360 ASFRILYSLENDIHKFRD 375

RESULT 3

beta IG-M2 protein precursor - mouse
 A:Accession: A40578
 C:Species: Mus musculus (house mouse)
 C:Date: 06-Mar-1992 #sequence_revision 06-Mar-1992 #text_change 01-Dec-2000
 C:Accession: A40578; A53228

R:Brummer, A.; Chinn, J.; Neuberger, M.; Purchio, A.F.
 DNA Cell Biol. 10, 293-300, 1991
 A:Title: Identification of a gene family regulated by transforming growth factor-beta.
 A:Reference number: A40578; MID:91229699; PMID:2029337

A:Accession: A40578
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-348 <BRD>
 A:Cross-references: GB:M80263; NID:9201945; PIDN:AAA73135.1; PID:9201946
 R:Hysek, R.P.; MacDonald-Bravo, H.; Mattei, M.G.; Bravo, R.
 Cell Growth Differ. 2, 225-233, 1991
 A:Title: Structure, mapping, and expression of fisp-12, a growth factor-inducible gene.
 A:Reference number: A53228; MID:91363290; PMID:1888698
 A:Accession: A53228
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-160; 'K', 162-348 <RYS>
 A:Cross-references: GB:M70641; NID:9193113; PIDN:AAA37627.1; PID:9193314
 C:Genetics:
 A:Gene: fisp-12

Query Match 45.5%; Score 957.5; DB 2; Length 348;
 Best Local Similarity 46.8%; Pred. No. 1.4e-60;
 Matches 178; Conservative 60; Mismatches 103; Indels 39; Gaps 8;

QY 1 MSSSTRFLAVATLHTL-TRLAL-STCPAACHPLE-APKCAPGVGLVVDGCGCCRYCA 57
 ||:|:|||||
 Db 1 MLASVAGPISLALVILALCTRRATGDCSAQCQCAEAAPHPACVSLVLDGCGCCRYCA 60
 QY 58 KOLNEDCSKTPQCDHTKGLCECNFGASSTALKGICRAQSEGRPCENSRIVONGESFOQPN 117
 |||||
 Db 61 KOLGELCTERDPCDPHKGJFCDFGSPANKRIGVCTAK-DGAPCVFGGSVYRSGESFOQSSC 119
 QY 118 KHQCTCIDGAVGCTPLCPQELSLPNCPCPNRLVYKVSGCCCEWVCEDESIKOSLDQDDL 177
 |||||
 Db 120 KHQCTCIDGAVGCTPLCPQELSLPNCPCPNRLVYKVSGCCCEWVCEDESIKOSLDQDDL 168
 QY 178 LGLDASEVELTNNNELIAGKSSSLKRL-PVEGTEPRVLPFLAHGOKICVOTTSNOCSSQ 236
 ||:|:|||||
 Db 169 -----KDRTRAVGPAALAYRLIEDTFEGDPTMM-----RANCLVOTTEMSA 207
 QY 237 CSKSGCTGISTRTVNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKCSKTKKSPPEVFT 296
 |||||
 Db 208 CSKSGCTGISTRTVNDNPECLVKEIRICEVRPCGQPVYSSLLKGGKCSKTKKSPPEVFT 267
 QY 297 TYAGSSVKKYRRKYGCSYDGRCTPLQTRVTKMRFRCDGEMFSKNVMIOCKKNYCN 356
 |||||
 Db 268 ELISGCTSVYTRAKFCGVCTDGRCTPHRTTLPEFKCPDEIMKMMFIKTCACHYN 327
 QY 357 CPHPNEASFRLYSLENDI 374
 ||:|:|||||
 Db 328 CPHPNEASFRLYSLENDI 347

RESULT 4

connective tissue growth factor - human
 A:Accession: A40551
 C:Species: Homo sapiens (man)
 C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 21-Jul-2000
 C:Accession: A40551; S44205

R:Bradham, D.M.; Igarashi, A.; Potter, R.L.; Grotenhorst, G.R.
 J. Cell Biol. 114, 1285-1294, 1991
 A:Title: Connective tissue growth factor: a cysteine-rich mitogen secreted by human v

A:Reference number: A40551; MID:91373462; PMID:1654338
 A:Accession: A40551

A:Molecule type: mRNA
 A:Residues: 1-349 <BRAS>

A:Cross-references: GB:M92934; GB:M36965; GB:S56201; NID:9180923; PIDN:AAA91279.1; PI

submitted to the EMBL Data Library, April 1994
 A:Description: Differential cloning and expression of human connective tissue growth

A:Reference number: S44205
 A:Accession: S44205

A:Status: preliminary
 A:Molecule type: mRNA

A:Residues: 1-349 <OEM>
 A:Cross-references: EMBL:X78947; NID:9474933; PID:9474934

[illegible]

```

RESULT 5
S20078
NOV protein - chicken
C:Species: Gallus gallus (chicken)
C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000
C:Accession: S20078
R:Joliot, V.; Martinet, C.; Dambrine, G.; Plasiart, G.; Brisac, M.; Crochet, J.; Pertin, M.
Mol. Cell. Biol. 12, 10-21, 1992
A:Title: Proviral rearrangements and overexpression of a new cellular gene (nov) in myeloid leukemia
A:Reference number: S20078; MUID:92107157; PMID:1309586
A:Accession: S20078
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-351 <JOL>
A:Cross-references: EMBL:X59284; NID:963702; PIDN:CAA41975.1; PID:963703
C:Genetics:
A:Gene: NOV

Query Match          40.9%; Score 861; DB 2; Length 351;
Best Local Similarity 44.5%; Pred. No. 9.6e-54;
Matches 165; Conservative 44; Mismatches 110; Indels 52; Gaps 7;

QY  9  LAVAVTLHLHTRLA-----LSTCPAC--HCPLEAPKAPGVGLVRDGGCGCCKCAQOLN 61
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  9  LPVLLLLLLLLLPREVSGREACPRPGGRCRPAAPPCAPVPAVLDDGGGLCLCANQRG 68

QY  62  EDCSKTOPCDHTKGLKCNFGASSTALKGICRAOSEGRPCENSRITYONGESFQPNCKHQC 121
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  69  ESCSPLLPCESSGGGLYCDRGPEDGGAGICNV-LEGNCCVFDGMIVYNGETFPQSCRYOC 127

QY  122 TCIGAVGCGPLCGQELSLPMLGCPNPRIVKVSGCCCEEWCD-EDSI-----KDSLDQ 175
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  128 TCRDQIGICLPKMLGILLPEPDCPPFRKLTVEPECEKWCDCPRDEVLGGFAMAAYRQ 187

QY  176 DDLGLADSEVELTRNNELIAIGKSSLRLLPVGTGPRVLFNPLHAHGOKCIYQTSMS 235
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  188 EATLGIVSD-----SSANCIQPTIEMS 210

QY  236 QCSKSCGTGISTRVNDNPBCRLVKETRICVEVRPGCGPVYSLKKKKCSKTKKSPEDVR 295

```

Db	211	ACSKCGMGFSRYATNRNDGCEMVAQTRLCMMRPCENEBSPD-KKKCKKIQTKKSKAVR	269
Qy	296	FTYAGCCSSVKKRYPRYCGSCVDGRCTPLQTRPYKMFCEDEGEMFSKRVMMIQSCKCN	355
Db	270	FEYKNCSTQVQYKPPRYCGJCNDRCTPHNTKIQVEFCPOGKFLKCPMMLINTCVCHG	329
Qy	356	NCPHPNEASFR	366
Db	330	NCPOSNNAFQ	340

```

RESULT 6
I38069
gene novH protein - human
C:Species: Homo sapiens (man)
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 05-Nov-1999
C:Accession: I38069
R:Matthierle, C.; Huff, V.; Joubert, I.; Badzioch, M.; Saunders, G.; Strong, T.; Perle,
Oncogene 9, 2729-2732, 1994
A:Title: Structural analysis of the human nov proto-oncogene and expression in Wilms
A:Reference number: I38069; MUID:94336229; PMID:7520150
A:Accession: I38069
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-357 <RES>
A:Cross-references: EMBL:X78351; NID:g587422; PIDD:CA455146.1; PID:g825696
C:Genetics:
A:Gene: novH
A:Introns: 28/3; 104/1; 188/1; 259/3
C:Superfamily: thrombospondin type I repeat homology
A:203-250/Domains: thrombospondin type I repeat homology <THRI>

```

[illegible]

submitted to the EMBL Data Library, October 1998
A:Reference number: Z20293

A:Accession: T26972

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1111 <MIL>

A:Cross-references: EMBL:AL032657; PIDN:CAA21739.1; GSPDB:GN00019; CESP:Y47H9C.4

A:Experimental source: clone Y47H9C

C:Genetics:

A:Gene: CESP:Y47H9C.4

A:Map position: 1

A:Insertions: 50/2; 84/2; 150/1; 238/3; 342/3; 797/1; 851/1; 947/2; 1017/1; 1083/1

C:Superfamily: unassigned ankryrin repeat proteins; ankryrin repeat homology; EGF homology

Query Match 8.1%; Score 171; DB 2; Length 1111;

Best Local Similarity 20.1%; Pred. No. 0.00016;

Matches 86; Conservative 33; Mismatches 156; Indels 154; Gaps 20;

```

QY 26 CPACACHPLE-ARKCAGVGLV-----DGGCCCKVCAK-QLNEDSKTQPCDHTKGLK 78
DB 380 CSTCTCTVRENTLMCAVNTFCRCRKGPFYGDNCCLACSKDSYGPNCCKAMCMNNASEC 439
QY 79 NFCASTALGICRAOSEGRPCENSRITYNGSEFOPNCKHOCCT-----IDGA----- 127
DB 440 NPETGSC-----VCKPGRITGNKSEPCPL-----DFYGPNCANHCOCQCNORVGCDBADGCKQ 491
QY 128 -----VGCIPD-----CPQELSLPN--LGGP-- 146
DB 492 CDKGWTHGRCEHHCPADTFGANCEKRCRCKPGIGCDPIITGECPCPAGLQCANCDIGCPDG 551
QY 147 -----NPLRYKVSQGCCCEWCVDESDINDSDDDDLGLDSEVELRNNE 193
DB 552 SYRPGCKLHCKCVNKGCKDKETGEC---TC-----OQGFEGSDS----- 587
QY 194 LIAIGKSSSLKRLPVETGEERVLFPNPLAHGOKCIVQ-TTWSQCSK---SCGTGISTRY 249
DB 588 -TTCSNCK-----YGESCELSGPCSDASCSQTKCLCPLSTKG 625
QY 250 TNDNPECRIVKTRICEVRPCGQPVYSSLRKGRKSKTKRSPPEVFTY-----AGC 301
DB 626 VSCDQCKDPMTPFGICETVTPSPCASTDPRKNGVCLSCPGSSGSIHCENCPAGSYDGC 685
QY 302 SSYKKYRPRKCGSCVGRCTPIQTRIV-----KMRRCEDGEMFSKNVMMIQSKC 353
DB 686 QQV-----C-SCADGHGDDPTTGECICEPGYHGKTSEKCPDGKYGICALDPCPKAS 737
QY 354 NYNCPHPN 361
DB 738 GSWCDHIN 745

```

RESULT 8

JC5598

mucin - rat

C:Species: Rattus norvegicus (Norway rat)

G:Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 05-Nov-1999

C:Accession: J05598

R:Inatomi, T.; Tisdale, A.S.; Zhan, Q.; Spurr-Michaud, S.; Gipson, I.K.

Biochem. Biophys. Res. Commun. 236, 789-797, 1997

A:Title: Cloning of rat Muc5AC mucin gene: Comparison of its structure and tissue distri-

A:Reference number: J05598; MUID:97396181; PMID:9245735

A:Accession: J05598

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-1034 <INA>

A:Cross-references: GB:083139; NID:g2315984; PIDN:AAC53312.1; PID:g2315985

C:Note: translation not complete

C:Comment: This protein is a high molecular weight glycoprotein which is a major compone

ntestinal tract and reproductive tract.

C:Genetics:

A:Gene: Muc5A

C:Superfamily: von Willebrand factor type C repeat homology

F:45-149/Domain: cysteine-rich <CTS>

F:762-830/Domain: von Willebrand factor type C repeat homology <VWC>

Query Match 7.9%; Score 166; DB 2; Length 1034;

Best Local Similarity 21.4%; Pred. No. 0.00055;

Matches 99; Conservative 47; Mismatches 170; Indels 146; Gaps 25;

```

QY 30 CHCPLEAPKCAPGVGLVRDGGCCCKVC---AKQINEDSKT-----OPCDHTKGLKCN 79
DB 558 CHM-LDLEVYCSGLIELYASLCAAGVCIPMRSHNTNTCPFTCPENVOYPCGSPNHYCY 616
QY 80 FGASSTALGICGA--QSEBRPCENSRITYO-----NESPDP----- 115
DB 617 RNDDISLALIQAGKRSCECFCDMTLFSSNDSICVSCQCLCPHPEVPEPGHTISI 676
QY 116 NCKHOCCTIDGAVGC-IPLCQPELSPLNLCPPNRLVKS-----GCCCEWVCDEDSIK 169
DB 677 NCQ-DCICKRGTITLCEKCLCPQT-----CPRGVPVPSIALENAGCCSQSFSC-----VC 725
QY 170 DSDDDDDDLGLDASEVELTRNNELIALIGK-SLKLRLPVFTGEPVLENP-----LHANG 224
DB 726 NSHCHCPPPLHCPRESSSLIVYERGTCCPSQNCSSQKGCVDVNGT---LYOPGDVYSSSLC 781
QY 225 QKCIYQVTS-----WSQCSKSGTGISRYVT-----NPNPFC 256
DB 782 ERLICEVSSNAFSDVAVVNCETELCMTQCPKGFETOTTEGHCCGCGVCPATCPKKNNST 841
QY 257 RLVEK-----TRICE-----VR-----PCGQP-- 279
DB 842 SLYKPGFEMPEPNPCVTHNCEKFDVLFVVTWKIECPKINCPDMAQLREDGCDYCLV 901
QY 280 KKKKSKTKRSPPEVFTTAAGCSSVKKRYKTC-GSCVDC-----RCQIP 323
DB 902 POKCKVTHQK-QILR-QQNCSESPVSLSYQGNCSGOSTSMYSLEANTVEHTCECQE 958
QY 324 LQRTVKMRPCEDGEMFSKNVMMIQSKC-NYNCPHPN 364
DB 959 LQTSQSRVTLHCDGSSRFSTYQVEKCGCLGQRCHAPDTS 1000

```

RESULT 9

T03099

mucin, submaxillary - pig

N:Alternate names: apomucin

C:Species: Sus scrofa domestica (domestic pig)

G:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 02-Jun-2000

C:Accession: T03099; A40009; A28528; B29789

R:Reckhardt, A.E.; Timpe, C.S.; Deluca, A.W.; Hill, R.L.

J. Biol. Chem. 272, 33204-33210, 1997

A:Title: The complete cDNA sequence and structural polymorphism of the polypeptide ch

A:Reference number: Z14839; MUID:98070526; PMID:9407109

A:Accession: T03099

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1328 <EBC>

A:Cross-references: EMBL:AF005273; NID:92501863; PIDN:AA62527.1; PID:g2581664

R:Reckhardt, A.E.; Timpe, C.S.; Abernethy, J.L.; Zhao, Y.; Hill, R.L.

J. Biol. Chem. 266, 9678-9686, 1991

A:Title: Porcine submaxillary mucin contains a cysteine-rich, carboxyl-terminal domain

A:Reference number: A40009; MUID:91236743; PMID:2033060

A:Accession: A40009

A:Molecule type: mRNA

A:Residues: 12139-12167, 'r', 12169-1328 <EC3>

A:Cross-references: GB:M61883; NID:9454837; PIDN:AAA30990.1; PID:g164374

R:Timpe, C.S.; Reckhardt, A.E.; Abernethy, J.L.; Hill, R.L.

J. Biol. Chem. 263, 1081-1088, 1988

A:Title: Porcine submaxillary gland apomucin contains tandemly repeated, identical se

A:Reference number: A28528; MUID:88087170; PMID:2826455

A:Accession: A28528

A:Molecule type: mRNA

A:Residues: 12139-12167, 'r', 12169-12641 <TIM>

A:Cross-references: GB:M21174; GB:J03512; NID:g164321; PIDN:AAA30990.1; PID:g552360

A:Experimental source: submaxillary gland

R:Reckhardt, A.E.; Timpe, C.S.; Abernethy, J.L.; Tounadje, A.; Johnson Jr., W.C.; Hill

J. Biol. Chem. 262, 11339-11344, 1987
 A:Title: Structural properties of porcine submaxillary gland apomucin.
 A:Reference number: A92606; MUID:87280230; PMID:3611111
 A:Accession: B29789
 A:Molecule type: Protein
 A:Residues: 1572-1607 <EC2>
 C:Superfamily: pig submaxillary mucin
 C:Keywords: tandem repeat

Query Match 7.5%; Score 157.5; DB 2; Length 13288;
 Best Local Similarity 20.1%; Pred. No. 0.013;
 Matches 93; Conservative 46; Mismatches 148; Indels 175; Gaps 22;

OY 26 CPACHCPLEAPKCAPGVGLVRDGGCGCKVCAKQLENECSKTQPCDHTKGLNECNAGAST 85
 DB 862 CKRCQYCP-----VGVNRNSKNC-----VFEDCC----- 886
 OY 86 ALKGICRAOSEGRPCENSRITYONGESFQPNCKHOCTCIDGAVGCP----- 132
 DB 887 -----PCSFGRREYDOGSVTSVGC-NKCTCIKSNCTONECQTTCHITGEG 932
 OY 133 -----LCPQELSLPNIGCPNP--RLVKVSGGCCCEWVDCDEDSIKSLDDQD 176
 DB 933 HIRFEDGKTYSFQGLQYSLFEDYCGSENGTFRILTESVPCCEDELTCRSKRIIVAFODON 992
 OY 177 DLL-----GLDASEVELRNNEEL-AIGKSSLRPLVF-----GTEPRVLEFNP 220
 DB 993 VLIHDKVTAVKTESKECELEENSVHTVGLYLILFLSGITIIIMDKNRISVILDP- 1051
 OY 221 HAHGOKC-----IVQTTSMSCSKSGTGISTRTVNDNPE 255
 DB 1052 RMNGKVCGLCGNNNDLDDFTTRSSAVAGELGNGSNKTSQESDVTGSPDSDNPY 1111
 OY 256 CRVKEIRICEV-----RPGQOPYSS-----LKKGKCSKTKKSPVPVFTYAG--CS 302
 DB 1112 CK-AWAVRKCETIIRDSTFRDCHNKVDPSAYDACEAEACDMCK-----YLGFC 1162
 OY 303 SVKKY-----RRKYGSCVD-----GRCC--TPLOTRTVKMFRCEDEM 340
 DB 1163 AWAAMAEKASAVGCVTRKDPCLCPVCDYNNAPBECRWREPCCTYAK--TCCKRVI 1219
 OY 341 FSKNYMMIOSC--KCNYNCPHPNEASFRYLSL-----FNFI 374
 DB 1220 GQKFSALLEGCYAKCPDSAPYIDEMTMKCVSLSECSCTYNDI 1261

RESULT 10

T13954

MEGF6 protein - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000

C:Accession: T13954

R.Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.

Genomics 51, 27-34, 1998

A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs

A:Reference number: Z14126; MUID:98360089; PMID:9693030

A:Accession: T13954

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1574 <NAK>

A:Cross-references: EMBL:AB011532; NID:93449293; PIDN:BA43462.1; PID:93449294

A:Experimental source: strain Sprague-Dawley; brain

C:Genetics:

A:Gene: MEGF6

Query Match

Best Local Similarity 7.4%; Score 156.5; DB 2; Length 1574;
 Pred. No. 0.0024;
 Matches 94; Conservative 40; Mismatches 142; Indels 185; Gaps 24;

OY 26 CPACHCP-----LEAPKCAPGV--GLVRDGC----- 50
 DB 568 CSSPCTCQNGTCDPVLGACRCRCPGVSAHCEGCPKGFYGHKCRKCHCANRGCHRLY 627

OY 51 GCKRVCAKOL-----NEDCSKTQPCDHTKGLNECNAGASTALKGICR 92
 DB 628 GAC-LCDBGLIGRFLHACLPFAEPGSGEDLCQ--SHRSCNPKGSGS-----CK 678
 OY 93 AOSEGRPCENSRITYONGESFQPNCKHOCTCIDGAVGCP-----CPQ 136
 DB 679 AGFQGERQAE-----ESG-FRQPGCRHHTCTQPG--VACDPVSGECRTQCPQYGEDCGQ 733
 OY 137 ELISLPNLG-----CENPRLVKYSQC--CEWVDCDEDSIKSLDDDDLLGLDASEVE 187
 DB 734 ECPVGTFFGVNCSGSCSGVAGPCHRVLTGBCLCPGKTGDBC--GADCEGRWGLCOET- 789
 OY 188 LTRNNELIAGKSSLRPLVPVTEPRVLEFNPVLAHAGKCIQVOTS-----WSQSK 239
 DB 790 -----CPACHGASCPNETGTCTCLPQFVSGKOD 819
 OY 240 SC-----GTGISTR--VTNDN-----PCRLVKEIRICE-----VRPC-- 270
 DB 820 TCSAGWGTGQIRCAACANDGHCDPTTGRCSCAPGMTGLSCGRACDSGHMGPDCIHPNC 879
 OY 271 -----GQPYSSLKKGKCKSKTKKSPVPVFTYAGSSVKKTRPKYCGSC--VD 317
 DB 880 SAGHCNDVAGSLCLCEAGYEGPRCEQS-----CROGYYPGSCDQKCRCEHGAACDHY 933
 OY 318 GRCCPTLOTRTYKMFRCDEDEMFESKNYMMIOSCKCNCP 358
 DB 934 GACTCPAGWRGSPCEHACPAGEF--GLDCDSACNCSAGAP 971

RESULT 11

T27283

hypothetical protein Y64G10A.f - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27283

R.Ainscough, R.

submitted to the EMBL Data Library, September 1999

A:Reference number: Z20336

A:Accession: T27283

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1620 <WIL>

A:Cross-references: EMBL:AL110498; NID:e1542303; PIDN:CA54471.1; CESP:Y64G10A.f

A:Experimental source: clone Y64G10A

C:Genetics:

A:Gene: CESP:Y64G10A.f

A:Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 548/1; 559/1; 601/1; 625/1

Query Match

Best Local Similarity 7.4%; Score 156.5; DB 2; Length 1620;
 Pred. No. 0.0024;
 Matches 92; Conservative 27; Mismatches 134; Indels 143; Gaps 22;

OY 26 CPACHCPLEAP-----KCAPG-----VLVRDGGCGCKVCAKQLENECSK 66
 DB 1052 CKGICSCQNGATCDSDVTSSCECRPMRKKDKDRCPDG--RFGECNMICTCTTNDISM 1109
 OY 67 TOP-----CDHTKGLNECNFASSTALKGICRAOSEGRPCENSRITYONGESFQPNCKHOCT 122
 DB 1110 YNPFVARCDHYTG--ECR-----CPAGMTGPDCQTSCLPGRHGE--GCRHSCQ 1153
 OY 123 CIDGA-----VGCIPCLPQELSLPNIGCPNPRLVKYSGCCCEWVDCDE 165
 DB 1154 CSNGASCDRVTFCDCCPSGFMKKNCESECPGLWGSN--CMKHCLCMHGGECNKE----- 1206
 OY 166 DSIKSLDDODDLGLDASEVELTRNNELIAGKSSLRPLVPVTEPRVLEFNPVLAHAG 225
 DB 1207 -----NEDCSCTQPCDHTKGLNECNAGASTALKGICR 92
 OY 226 KCIYOT-----TSMS--QCSKSGCTGISTRTVNDNPRCLVKEIRICEVPRCPQPYSS 277
 DB 1242 SCDRTGRCLELPGMSGHCHCKSCVSG-----HYGAKC--EETCEBENGLADPISGH 1292
 OY 278 LK-----KGKCSKTKKSPVPVFTYAGSSVKKTRPKYCGSC--VDGR----- 320

Db 1293 CSCOPGWRGRKKNR-----PCLKGYFGRHCOSQSCRCANSKSCDHISGRCOCPEKGYAGHS 1346
 QY 321 CTPLQTRTVKMRFCEDGEMFSKVMVMIOOSCKMYN 356
 Db 1347 CTCL-----CPDG-TFGSCS--OKCDDGEN 1369

RESULT 12

T42626

secreted leucine-rich repeat-containing protein SLIT2 - mouse (fragment)

N:Alternate names: neurogenic extracellular slit protein

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 16-Aug-2002

C:Accession: T42626

R:Holmes, G.P.; Negus, K.; Burridge, L.; Raman, S.; Alagar, E.; Yamada, T.; Little, M.H.

Mech. Dev. 79, 57-72, 1998

A:Title: Distinct but overlapping expression patterns of two vertebrate slit homologs in

A:Reference number: 422177; MUID:9279238; PMID:10349621

A:Accession: T42626

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1025 <HOL>

A:Cross-references: EMBL:AF074960; NID:g4151258; PID:g4151259; PIDN:AD04345.1

C:Genetics:

A:Gene: Slit2

C:Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein

Query Match

Best Local Similarity 7.4%; Score 156; DB 2; Length 1025;

Matches 101; Conservative 41; Mismatches 150; Indels 248; Gaps 23;

QY 24 STC-----PACHCPLPAPKCAPGVGLVRDGGCGCKYCAQOLNEDCSKTOPCDH----- 72
 Db 507 STCVGDGINNYTCLCPPEYTG-----ELCEKELDFCAQDLNPGQIDSKCIL 551
 QY 73 -TGLECNFGASSALKGICRAOSEGRPCYNSRIYNGSFQPNCHQCTCIDGAVCI 131
 Db 552 TPFGFKD-----CTPGYIGHCIDIDDDCDNK-----CKNGAHCDDAVNGYT 595
 QY 132 PLCPQ-----ELSLP-----NLGCPNPR--LVKYS-----GOCCEW 161
 Db 596 CVCPEGSGLECFESPWVLPRTSPCNFPCQNAQCIIRINEICQCLGYLEKEKEL 655
 QY 162 VCEDESLIKDS-----LDDQDDLGL-----DASEVELTRNNELIAI 197
 Db 656 VSNFVFKESYLQIPSAKVRPQNTITLQIATDEDSGLLYKGDKDHIAVELYRGVASY 715
 QY 198 GKRS-----SLKRLPYGTEPRVLEN-----PL 220
 Db 716 DTGSHPASATISVETINDGNFIVEILLTLDLSLSLVYDGGSPRYITNLSKOSTLNEDSPL 775
 QY 221 HA-----HG-----OKCIVQ 230
 Db 776 YVGMPEKKNVASTLRAPGQNGTSFHCIRNLINSELQDFRKMPQGTGLPGCEPHKK 835
 QY 231 TTSMQCSKSGTGISTRV-----TND----- 252
 Db 836 VCAHGMQPSQSGFTCEEGEMWGPLCDORTNDPCLGKNCVHGTCPLINAFYSCKCLE 895
 QY 253 -----NPECRIVK-ETRLCEVRPGQPV--YSLKKGKCKK-TKSKPEPV 294
 Db 896 GHGGLVLCDEEDELFPN-CQMIKCKHGKCRISGVGPCECNSGFTGSDSCREISCRERI 954
 QY 295 R-----FTYAGCSSVKKYRKYC-GSCVDGRCTPLQTRTVKMRFCEDGEMFSKVM 347
 Db 955 RQYVQKQGGYAAQOTTKKVRSLRLECRGCGAGGCCGGLRSKRRTYSPECTGSSSFVDEYK 1014
 QY 348 IQSCCK 353
 Db 1015 VVKCGC 1020

RESULT 13

A57534

mucin 5AC (clone L31) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 20-Apr-2000

C:Accession: A57534

R:Lesuifleur, T.; Roche, F.; Hill, A.S.; Lacasa, M.; Fox, M.; Swallow, D.M.; Zweibaum

J. Biol. Chem. 270, 13665-13673, 1995

A:Title: Characterization of a mucin cDNA clone isolated from HT-29 mucous-secreting c

A:Reference number: A57534; MUID:95293957; PMID:7775418

A:Accession: A57534

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1042 <LES>

A:Cross-references: GB:448314; NID:g1052607; PIDN:CAA88307.1; PID:g1052608

A:Gene: GDB:MUC5AC

A:Cross-references: GDB:454136; OMIM:158373

A:Map position: 11p15.5-1p15.5

A:Superfamily: von Willebrand factor type C repeat homology

F:678-746/Domain: von Willebrand factor type C repeat homology <VWC>

Query Match 7.3%; Score 152.5; DB 2; Length 1042;

Best Local Similarity 18.9%; Pred. NO. 0.0037; Indels 245; Gaps 27;

Matches 105; Conservative 48; Mismatches 159;

QY 4 STERTIAVAATLHLRLALSTCPACCHPLAPKCAPGVGLVRDGC-----GC 52
 Db 411 TTVGSGTVGPTTVGSGTVGPTTPPAPC---LPSPICHLISLVFEBCHVYIPPLFYECC 467
 QY 53 -----CKYCAK-----QLNECSKT-----QCCHITKGL 76
 Db 468 VFDRCHMTLDLVYSSLELYAALCASHDIDNRGRTGHMCPTCPADKRYVQCPGSPNS 527
 QY 77 ECNFGASSTALKGICRAQ--SEGRPCYNSRIYON-----GE 111
 Db 528 YC-YGNDASLGLAPRAGPTTEGCFPEBGITLSTNAQCVPTGRCRLGHPHYKGVH 586
 QY 112 SFQPNCHQCTC--IDAVGCIPL-ICQELSLPNLCCPNRLVYV-----SGQCEEW 163
 Db 587 TVGMDQ-ECTCPAATWTILCRPKLCP---LPP-ACPLPGFVVPVAPAPQAGCCPQYSC 640
 QY 164 DEDSINDLDDQDDLGLDASEVELTRNNELIAGGSSILKRLPV-FGEPRVLENPLH 222
 Db 641 -----ACNSTRCPAVGCEGARALPTQ 664
 QY 223 HGQCIYQTSWQCSKSGCTGISTRYTNDNPECRIVK-----ETRICE 266
 Db 665 EGACCPYQNCSTWVCSIN-GTLQPGAVYSSISCEICRCLPGGPPSDAFVYSCETQICN 723
 QY 267 V-----RPGQPV--YSLKKG-----KCKSK-- 286
 Db 724 THCPVGFEYQESGQCGCTGVQAVCTNTSKSAHLFPGETWSDAGNHCVTHQCEKHOD 783
 QY 287 -----TKSKPEPV-----RF-----TYAGCS 302
 Db 784 GLVYVTTKRCAPLSLSLEARMKSKDCCRFCLPPPPYONOSTCAVYHNSLIIDQOGCS 843
 QY 303 SVAKYRKYC-GSCVCG-----RCQTPLOTRTVKKMRFCEDGEMFSKVM 346
 Db 844 SSEPVRILANCRCGCGSSSYSLGNTVEHRQCCQELRLSLNVLHCTDGGSSRAFSIT 903
 QY 347 MIOCKC-NYNCPHNE 362
 Db 904 EVEECGCMGRCPAPGD 920

RESULT 14

A53767

mucin MUC5B, tracheobronchial - human

C:Species: Homo sapiens (man)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Apr-2000

C:Accession: A53767

R.Meerzaman, D.; Charles, P.; Daskal, E.; Polymeropoulos, M.H.; Martin, B.M.; Rose, M.C.
 J. Biol. Chem. 269, 12932-12939, 1994
 A:Title: Cloning and analysis of cDNA encoding a major airway glycoprotein, human trachea
 A:Reference number: A53767; MUID:94230376; PMID:7513696
 A:Accession: A53767
 A:Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1056 <ME>
 A:Cross-references: GB:006711; NID:9488072; PID:AA18431.1; PID:9488073
 A:Note: Authors translated the codon TAT for residue 1054 as Thr
 C:Genetics:
 A:Gene: GDB:MUC5B; MUC5
 A:Cross-references: GDB:128436; OMIM:600770
 A:Map position: 11p15.5-11p15.5
 C:Superfamily: von Willebrand factor type C repeat homology
 F:67-835/Domain: von Willebrand factor type C repeat homology <MC>

Query Match 7.2%; Score 152; DB 2; Length 1056;
 Best Local Similarity 19.1%; Pred. No. 0.0035;
 Matches 107; Conservative 57; Mismatches 152; Indels 244; Gaps 30;

```

QY 4 STEFLAVATLHLRLALSTCPACHPLPAKCAPGVLDGC-----GC 52
DB 501 TTGCTVGPPTVSTVGPPTPAPC---LPSPICHLILSKVEPCHTVIPPLTFYEGC 557
QY 53 -----CKVCAKQINEDC-----SKTOPCDH-----TKGLE----- 77
DB 558 VFDCCHMTDLDVYSSSLFLVRLCAS--HDICIDMRGTRICAHHLPSRQGVLPALRSPNP 615
QY 78 --CNFGASSTALKGICRAO--SEGRPCENSRIRYON-----G 110
DB 616 SYC-XGNDASALGLAREGPTTEGCFEGMTLESTAOVCVPTGCPRLGPHGPVAVG 674
QY 111 ESFPNCKHQTCT--IDGAVGCP--LCPQELSLPMLGCPNRLYKV-----SGQCEEMV 162
DB 675 HTVMDCC-ECTCEAATWTLTCRPLCP---LPP-ACPLPGFVPAAPAGOCOPQYS 728
QY 163 CDEDSINDSLDDQDGLGLDASEVELTFNNELIAGKSSSLKRLPVPTETERRVLEFNLHA 222
DB 729 C-----ACNTSRCPAPVCPGSA--RIPTY-----Q 753
QY 223 HGQCIYQTVSMSC-----SKSCGT-----GISTRVIND 252
DB 754 EGACCPQNCMTWCSINGTLXPQAVYSSSLCTCRCELPGSPDAFVSCETQICNT 813
QY 253 NPEGRLL-VKETRICEVRP--CGOPYSSSLKKG-----KCKSK--- 286
DB 814 HCPVRFETQEQRRSAVAVPCRPSPVTPAPAPTSSTIASWSDAGNHCVTHQCEKHODG 873
QY 287 -----TKSPRPV-----FTYAGC-----SS 303
DB 874 LVVYTTKACPLPCLSLDEARMSKDGCCRCPLPPPYONOSTCAVYHRSLLIIOQGGSS 933
QY 304 VKKTRPKYC-GSCVDG-----RCCTPLQTRTVKMRFRCEDEMEFSKNVMM 347
DB 934 SEPRLAYCNRNGCDSSSMYSLEGNTEVHRQCCQELSTLRNVTLIHCTDSSRAFSYTE 993
QY 348 IQSKC-NTNCPHNEASF 366
DB 994 VEEGCGMRCPAPAPSTR 1013

```

RESULT 15

S08167
 Balblant ring 3 protein - midge (Chironomus tentans)
 C:Species: Chironomus tentans
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Jul-2000
 C:Accession: S08167
 R:Paulsson, G.; Lendahl, U.; Gallit, J.; Ericsson, C.; Wieslander, L.
 J. Mol. Biol. 211, 331-349, 1990
 A:Title: The balblant ring 3 gene in Chironomus tentans has a diverged repetitive structure
 A:Reference number: S08167; MUID:90172404; PMID:1689777
 A:Accession: S08167

A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-1700 <PN>
 A:Cross-references: GB:X52263; NID:97057; PID:CAA36506.1; PID:97058
 C:Genetics:
 A:Gene: BR3
 A:Map position: 4
 C:Superfamily: unassigned Balblant ring proteins

Query Match 7.2%; Score 151.5; DB 2; Length 1700;
 Best Local Similarity 22.7%; Pred. No. 0.0057;
 Matches 83; Conservative 41; Mismatches 143; Indels 99; Gaps 23;

```

QY 30 CHCLPAKCAPGVGLVNDGG--C-CKVCAKQINEDSKTOPCDHTGLEGCNFGASSTA 86
DB 384 CIGPDADVCTAG---KERGSECECINREPEEGAKP-----LVWN---ENT 427
QY 87 LKICRAOSEGRPCENSRIRYONGSEFQP--NCKHQTCTIDGAVG-----IPLCPQE 137
DB 428 CKCYCPADKQMSPGGCGS-----GKSFNKLTC--QCEBDQASAKGLKRWNAIDCKCECQ 480
QY 138 LSLPMLGCPNRLV--KVSQCEEMVDEDSIKDSLDDQDGLGLDASEVE---LTR 190
DB 481 PGMPPEGGCKQOTWISDKCKCECSPTITQAPQILD-----LNTCECKCPVNMIAQ 530
QY 191 NNELIAGKSSSLKRLPVPTETERRVLEFNLHAHQKCIYQTVSMSC-----SKSCGTGI 245
DB 531 KEKCKSPQMTWDSKCLCECSTTPATCEBQKQWCGAC-----QCICPGDGKNG--- 579
QY 246 STRVYNDNPECRIVKETRICEVRP--CGOP--VYSSSLKKGKCKSKTSPKSPRYFVAGCS 303
DB 580 -NKKFFDPSCECK-----CKNNPTCTSPQWADADDCCKCPKQKQKQ-----GGCDG 627
QY 304 VKKTRPKYCGSCVDGRCTPLQTRTVKMRFRCEDEMEFSKNVMMIQSKCNYNCPHNEA 363
DB 628 GQKWNDRVC-SC---GCPVP-----RPDCTNGQIYN-----INTCAGCGGIDKPCSP 670
QY 364 SFRLYS 369
DB 671 KQIYIN 676

```

Search completed: August 5, 2003, 14:07:27
 Job time : 17.4592 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:02:16 ; Search time 9.475 Seconds
(without alignments)
1881.068 Million cell updates/sec

Title: US-09-495-448a-2

Perfect score: 2103
Sequence: 1 MSSSTFRTLAVALTLHLTR.....PNEASFRLYSLEFNDHKFRD 379

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2103	100.0	379	1	CYR6_MOUSE
2	2079	98.9	379	1	CYR6_MOUSE
3	1938	92.2	381	1	CYR6_HUMAN
4	1646.5	78.3	375	1	CE10_CHICK
5	960	45.6	347	1	CTGF_RAT
6	957.5	45.5	346	1	CTGF_MOUSE
7	950	45.2	349	1	CTGF_HUMAN
8	937	44.6	349	1	CTGF_BOVIN
9	919	43.7	351	1	CTGF_PIG
10	861	40.9	351	1	NOV_CHICK
11	856	40.7	353	1	NOV_COTTA
12	839.5	39.9	357	1	NOV_HUMAN
13	829.5	39.4	351	1	NOV_RAT
14	827	39.3	354	1	NOV_MOUSE
15	824.5	39.2	343	1	NOV_XENLA
16	526.5	25.0	251	1	CTGL_MOUSE
17	524	24.9	250	1	CTGL_HUMAN
18	492	23.4	250	1	CTGL_RAT
19	158	7.5	447	1	NRL1_MOUSE
20	154	7.3	456	1	NRL1_CHICK
21	154	7.3	3110	1	LM22_MOUSE
22	152.5	7.3	1233	1	LM22_HUMAN
23	151.5	7.2	1700	1	BAR3_MOUSE
24	148	7.0	450	1	NRL1_HUMAN
25	146	6.9	837	1	MUC1_RAT
26	145.5	6.9	2813	1	WVF_MOUSE
27	144.5	6.9	5703	1	WVF_HUMAN
28	144	6.8	4655	1	LRP2_MOUSE
29	143	6.8	1172	1	WVF_PIG
30	141.5	6.7	1282	1	LM22_HUMAN
31	141	6.7	1187	1	LM22_MOUSE
32	141	6.7	5376	1	LM22_HUMAN
33	140.5	6.7	2282	1	ZAN_MOUSE

ALIGNMENTS

RESULT 1	ID	CYR6_MOUSE	STANDARD	PRT	379 AA
AC	PI8406				
DT	01-NOV-1990 (Rel. 16, Created)				
DT	01-NOV-1990 (Rel. 16, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)				
DE	(Insulin-like growth factor-binding protein 10) (IGH1).				
GN	CYR61 OR IGFBP10.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
OX	NCBI_TaxID=10090;				
RP	SEQUENCE FROM N.A.				
RP	STRAIN=BALE/C; TISSUE=Fibroblast;				
RC	MEDLINE=90287146; PubMed=2355916;				
RA	O'Brien T.P., Yang G.P., Sanders L., Lau L.F.;				
RT	"Expression of cyr61, a growth factor-inducible immediate-early				
RT	gene.";				
RL	Mol. Cell. Biol. 10:3569-3577(1990).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=AJ; TISSUE=Embryonic fibroblast;				
RX	MEDLINE=91288203; PubMed=2062642;				
RA	Latinkic B.V., O'Brien T.P., Lau L.F.;				
RT	"Promoter function and structure of the growth factor-inducible				
RL	immediate early gene cyr61.";				
RL	Nucleic Acids Res. 19:3261-3267(1991).				
CC	-1- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING				
CC	PROTEINS; PROMOTES PROLIFERATION, MIGRATION AND ADHESION.				
CC	-1- SUBCELLULAR LOCATION: Secreted.				
CC	-1- TISSUE SPECIFICITY: LOW IN KIDNEY, ADRENAL GLAND, TESTES, BRAIN,				
CC	AND OVARY, MODERATE IN HEART, UTERUS, AND SKELETAL MUSCLE, HIGHEST				
CC	IN LUNG.				
CC	-1- DEVELOPMENTAL STAGE: EXPRESSED FROM G(0)/G(1) THROUGH MID-G(1) IN				
CC	NORMAL CELLS, AND AT A CONSTANT LEVEL IN RAPIDLY GROWING CELLS.				
CC	-1- INDUCTION: By growth factors.				
CC	-1- SIMILARITY: Contains 1 IGF1R domain.				
CC	-1- SIMILARITY: Contains 1 WPC domain.				
CC	-1- SIMILARITY: Contains 1 TSP type-1 domain.				
CC	-1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	EMBL: M32490; AAA37512.1; -				
DR	EMBL: X56790; CAA40109.1; -				
DR	PIR: A35669; A35669.				
DR	MGD: MGI:86613; Cyr61.				

DR GO: 0001569; P: patterning of blood vessels; IMP.
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Insl_gro_fac_pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00219; IGFBP; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR Pfam: PF00093; vwc; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00209; TSP1; 1.
 DR SMART: SM00214; VMC; 1.
 DR PROSITE: PS01185; CTCK_1; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR PROSITE: PS00222; IGF_BINDING; 1.
 DR PROSITE: PS50092; TSP1; 1.
 DR PROSITE: PS01208; VWF_C_1; 1.
 DR PROSITE: PS50184; VWF_C_2; 1.
 DR PROSITE: PS50184; VWF_C_2; 1.
 KM Growth factor binding; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 379 POTENTIAL.
 FT DOMAIN 98 164 CYR61 PROTEIN.
 FT DOMAIN 226 271 VWF.
 FT DOMAIN 284 358 TSP TYPE-1.
 FT DISULFID 284 321 CTCK.
 FT DISULFID 301 335 BY SIMILARITY.
 FT DISULFID 312 351 BY SIMILARITY.
 FT DISULFID 315 353 BY SIMILARITY.
 FT DISULFID 320 357 BY SIMILARITY.
 SQ SEQUENCE 379 AA; 41709 MW; FA6B5014B56A8EB9 CRC64;

Query Match 100.0%; Score 2103; DB 1; Length 379;
 Best Local Similarity 100.0%; Pred. No. 5,6e-146;
 Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSSTFRTLA VAVTLTLRLALSTCPAACHPELPKAPGVGLVVDGCGCKVCANOL 60
 1 MSSSTFRTLA VAVTLTLRLALSTCPAACHPELPKAPGVGLVVDGCGCKVCANOL 60
 DB 1 MSSSTFRTLA VAVTLTLRLALSTCPAACHPELPKAPGVGLVVDGCGCKVCANOL 60
 DB 61 NECCSKTOPCDHRTKGLKCNFGASSSTALKGICRAQSBGRPEVNSRYIQNESTQPNCKHQ 120
 61 NECCSKTOPCDHRTKGLKCNFGASSSTALKGICRAQSBGRPEVNSRYIQNESTQPNCKHQ 120
 DB 61 NECCSKTOPCDHRTKGLKCNFGASSSTALKGICRAQSBGRPEVNSRYIQNESTQPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPNLGCPRPLVYKSGGCCWVDEDSIKSLDDODDLG 180
 121 CTCIDGAVGCIPLCPQELSLPNLGCPRPLVYKSGGCCWVDEDSIKSLDDODDLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNLGCPRPLVYKSGGCCWVDEDSIKSLDDODDLG 180
 QY 181 LDASEVELTRNNELIAGKSSSLKRLPVFETEPRLVFNPLHAHGOKCIYOTTSMSQCSKS 240
 181 LDASEVELTRNNELIAGKSSSLKRLPVFETEPRLVFNPLHAHGOKCIYOTTSMSQCSKS 240
 DB 181 LDASEVELTRNNELIAGKSSSLKRLPVFETEPRLVFNPLHAHGOKCIYOTTSMSQCSKS 240
 QY 241 CGGIGISRTYNDNPEGLVKEETRICVRRPGQPYVSSLLKGGKCKSTKKSPEVRYTYAG 300
 241 CGGIGISRTYNDNPEGLVKEETRICVRRPGQPYVSSLLKGGKCKSTKKSPEVRYTYAG 300
 DB 241 CGGIGISRTYNDNPEGLVKEETRICVRRPGQPYVSSLLKGGKCKSTKKSPEVRYTYAG 300
 QY 301 CSSVKKRRPYKCGSSCYDGRCTPLQITRYVMRRFCDEGEAFSNVMMIIOGCKNYNCPHP 360
 301 CSSVKKRRPYKCGSSCYDGRCTPLQITRYVMRRFCDEGEAFSNVMMIIOGCKNYNCPHP 360
 DB 301 CSSVKKRRPYKCGSSCYDGRCTPLQITRYVMRRFCDEGEAFSNVMMIIOGCKNYNCPHP 360
 QY 361 NEASFRLYSLFNDIHKFRD 379
 361 NEASFRLYSLFNDIHKFRD 379
 DB 361 NEASFRLYSLFNDIHKFRD 379

RESULT 2
 CYR6_RAT
 ID CYR6_RAT STANDARD; PRT; 379 AA.
 AC 09ES72;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)
 DE (Insulin-like growth factor-binding protein 10).
 GN CYR61 OR IGFBP10.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20435857; PubMed=10852911;
 RA Albrecht C., von Der Kammer H., Mayhaus M., Klaudiny J., Schweizer M.,
 RA Nitsch R.M.;
 RT "Muscarinic acetylcholine receptors induce the expression of the
 RT immediate early growth regulatory gene CYR61.";
 RL J. Biol. Chem. 275:28929-28936(2000).
 CC -1- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
 CC PROTEINS; PROMOTES PROLIFERATION, MIGRATION AND ADHESION (BY
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 VWF domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: AF218568; AAC14964.1; -
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Insl_gro_fac_pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00029; IGFBP; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR Pfam: PF00093; vwc; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00209; TSP1; 1.
 DR SMART: SM00214; VMC; 1.
 DR PROSITE: PS01185; CTCK_1; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR PROSITE: PS00222; IGF_BINDING; 1.
 DR PROSITE: PS50092; TSP1; 1.
 DR PROSITE: PS01208; VWF_C_1; 1.
 DR PROSITE: PS50184; VWF_C_2; 1.
 KM Growth factor binding; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 379 POTENTIAL.
 FT DOMAIN 98 164 CYR61 PROTEIN.
 FT DOMAIN 226 271 IGFBP.
 FT DOMAIN 284 358 VWF.
 FT DOMAIN 284 321 TSP TYPE-1.
 FT DISULFID 284 321 CTCK.
 FT DISULFID 301 335 BY SIMILARITY.
 FT DISULFID 312 351 BY SIMILARITY.
 FT DISULFID 315 353 BY SIMILARITY.
 FT DISULFID 320 357 BY SIMILARITY.
 SQ SEQUENCE 379 AA; 41687 MW; 62BF0BBA4C5AFDE9 CRC64;

Query Match 98.9%; Score 2079; DB 1; Length 379;
 Best Local Similarity 98.7%; Pred. No. 3.3e-146;
 Matches 374; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MSSSTFRTLA VAVTLTLRLALSTCPAACHPELPKAPGVGLVVDGCGCKVCANOL 60
 1 MSSSTFRTLA VAVTLTLRLALSTCPAACHPELPKAPGVGLVVDGCGCKVCANOL 60

DB 1 MSSSTIKTILAVATLHLFRLALSTCPAACHCPLFAKCAPGVLYNDGCGCCAKOL 60
 QY 61 NECSKTPQCDHRTKGLJECNCGASTALGICRAOSEBPREYNSRITONESPQPNKHQ 120
 DB 61 NECSKTPQCDHRTKGLJECNCGASTALGICRAOSEBPREYNSRITONESPQPNKHQ 120
 QY 121 CTCIDGAVGICPLCPQELSLPNIGCNPRILVYKSGCCCEWVDEDSIKSLDDQDILLG 180
 DB 121 CTCIDGAVGICPLCPQELSLPNIGCNPRILVYKSGCCCEWVDEDSIKSLDDQDILLG 180
 QY 181 LDASEVELLRNNELIAGKSSSLKRLPVGTEPRVLYFNPLHAHQKCIQVOTSMSCSKS 240
 DB 181 FDASEVELLRNNELIAGKSSSLKRLPVGTEPRVLYFNPLHAHQKCIQVOTSMSCSKS 240
 QY 241 CGGISTRTVNDNDPECLVETRICFVRPCGOPYSSLSKKGKCKSTKSPPEVRYTAG 300
 DB 241 CGGISTRTVNDNDPECLVETRICFVRPCGOPYSSLSKKGKCKSTKSPPEVRYTAG 300
 QY 301 CSSVKYRPRYKSCGVDRCTPLQRTVYKMRFCEDGEMFSKNVMMIOCKCNYNCPHP 360
 DB 301 CSSVKYRPRYKSCGVDRCTPLQRTVYKMRFCEDGEMFSKNVMMIOCKCNYNCPHP 360
 QY 361 NEASFRLYSLFNDIHKFRD 379
 DB 361 NEASFRLYSLFNDIHKFRD 379

RESULT 3
 CYR6_HUMAN STANDARD; PRT; 381 AA.
 ID 000622; 014934; 043775; 09BZL7;
 AC 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE CTR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)
 DE (insulin-like growth factor-binding protein 10) (IGFBP10).
 GN CYR61 OR IGFBP10 OR GIG1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Albrecht C., von der Kammer H., Klaudiny J., Mayhaus M., Nitsch R.M.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=97280750; PubMed=9135077;
 RA Jay P., Berge-Leiranc J.L., Marsollier C., Mejean C., Tavlaux S.,
 RA Berta P.;
 RL "the human growth factor-inducible immediate early gene, CYR61, maps
 to chromosome 1p.11-1753-1757(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98197344; PubMed=9536281;
 RA Martinier C., Viegas-Pequignot E., Nguyen V.C., Perbal B.;
 RT "Chromosomal mapping and expression of the human cyr61 gene in tumour
 cells from the nervous system";
 RL Mol. Pathol. 50:310-316(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Placenta;
 RA Kolesnikova T.V., Lau L.F.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA BI A.B., Yu L.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Schuetze N., Lechner A., Groll C., Koehrl J., Jakob F.;
 RT "Regulation of hcyr61 by vitamin D, serum and cytokines in fetal human

RT osteoblasts."; to the EMBL/GenBank/DBJ databases.
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Leng E., Tai G., Estable M., Liu J., Chow C., Sadowski I.;
 RT "Organization and expression of the cyr61 gene in normal human
 RT fibroblasts."; to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung, Placenta, and Skin;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Shat N.K.,
 RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Maruska K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshilyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wotley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
 CC PROTEINS; PROMOTES PROLIFERATION, MIGRATION AND ADHESION (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 WFC domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation at
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 DR EMBL: Y12084; CAA72802.1; -;
 DR EMBL: U62015; AAB58319.1; -;
 DR EMBL: Y11307; CAA72167.1; -;
 DR EMBL: AF003594; AAB61240.1; -;
 DR EMBL: AF031385; AAB84227.1; -;
 DR EMBL: Z98053; CAB10848.1; -;
 DR EMBL: AF307860; AAG59863.1; -;
 DR EMBL: BC001371; AA01371.1; -;
 DR EMBL: BC009199; AA009199.1; -;
 DR EMBL: BC016952; AA016952.1; -;
 DR Genew; HGNC:2654; CYR61.
 DR MIM: 602369; -;
 DR GO: GO:0008283; P:cell proliferation; TAS.
 DR GO: GO:0007345; P:embryogenesis and morphogenesis; TAS.
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Ins1_knot.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; WFC_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00219; IGFBP; 1.
 DR Pfam: PF00090; TSP_1; 1.
 DR Pfam: PF00093; WFC; 1.
 DR SMART; SM00041; CT; 1.

DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF_BINDING; 1.
 DR PROSITE; PS50092; TSP1; 1.
 DR PROSITE; PS01208; VWF; 1; 1.
 DR PROSITE; PS50184; VWF_2; 1.
 KM Growth factor binding; Signal.
 FT SIGNAL; 1 24
 FT CHAIN; 25 381
 FT DOMAIN; 26 97
 FT DOMAIN; 98 164
 FT DOMAIN; 228 273
 FT DOMAIN; 286 360
 FT DISULFID; 286 323
 FT DISULFID; 303 337
 FT DISULFID; 314 353
 FT DISULFID; 317 355
 FT DISULFID; 322 359
 FT CONFLICT; 165 165
 FT CONFLICT; 210 210
 FT CONFLICT; 220 220
 FT CONFLICT; 369 369
 SQ SEQUENCE 381 AA; 42026 MW; FC0BD39C078CA0B1 CRC64;

Query Match 92.2%; Score 1938; DB 1; Length 381;
 Best Local Similarity 91.4%; Pred. No. 8e-136;
 Matches 350; Conservative 9; Mismatches 18; Indels 6; Gaps 2;

QY 1 MSSSTRTLAVALVTLHLTRLALSTCPACCHCPLEAPKCAPGVGLVDDGCGCCVCAKOL 60
 1 MSSRIARALALVTLHLTRLALSTCPACCHCPLEAPKCAPGVGLVDDGCGCCVCAKOL 60
 DB 61 NECCSTQPCDHTKGLKCNFGASSTALKGICRAQSGRCCEVNSRTIYONGESQPNCKIQ 120
 61 NECCSTQPCDHTKGLKCNFGASSTALKGICRAQSGRCCEVNSRTIYONGESQPNCKIQ 120
 QY 121 CXCIDAVAGCIPCLPOELSLPNLGNCPNRLVKVSGGCEWVDEDSIKDSLDDODL-- 178
 121 CXCIDAVAGCIPCLPOELSLPNLGNCPNRLVKVSGGCEWVDEDSIKDSLDDODL-- 178
 DB 121 CXCIDAVAGCIPCLPOELSLPNLGNCPNRLVKVSGGCEWVDEDSIKDSLDDODL-- 180
 121 CXCIDAVAGCIPCLPOELSLPNLGNCPNRLVKVSGGCEWVDEDSIKDSLDDODL-- 180
 QY 179 --IGLDASEVELTRNNELIAGKSSILKRLPVFGTEPRVLFNPLAHAGOKCIYQVTSWQ 236
 179 --IGLDASEVELTRNNELIAGKSSILKRLPVFGTEPRVLFNPLAHAGOKCIYQVTSWQ 236
 DB 181 KELGFDASEVELTRNNELIAGKSSILKRLPVFGTEPRVLFNPLAHAGOKCIYQVTSWQ 238
 181 KELGFDASEVELTRNNELIAGKSSILKRLPVFGTEPRVLFNPLAHAGOKCIYQVTSWQ 238
 QY 237 CCKSCGTGISTRTNPNPCRLVKEIRICEVRCCGPVYSSLLKKGKCKSKTKSPPEVRF 296
 237 CCKSCGTGISTRTNPNPCRLVKEIRICEVRCCGPVYSSLLKKGKCKSKTKSPPEVRF 296
 DB 239 CSKTCGTGISTRTNPNPCRLVKEIRICEVRCCGPVYSSLLKKGKCKSKTKSPPEVRF 298
 239 CSKTCGTGISTRTNPNPCRLVKEIRICEVRCCGPVYSSLLKKGKCKSKTKSPPEVRF 298
 QY 297 TYAGCSVKKRYKRYGCGVDCGRCCTPLQTRPKMFRCEDEGMFSKNYMMIOSCCNCYN 356
 297 TYAGCSVKKRYKRYGCGVDCGRCCTPLQTRPKMFRCEDEGMFSKNYMMIOSCCNCYN 356
 DB 299 TYAGCSVKKRYKRYGCGVDCGRCCTPLQTRPKMFRCEDEGMFSKNYMMIOSCCNCYN 358
 299 TYAGCSVKKRYKRYGCGVDCGRCCTPLQTRPKMFRCEDEGMFSKNYMMIOSCCNCYN 358
 QY 357 CPHNPASFRFLYSLENDIHKFRD 379
 357 CPHNPASFRFLYSLENDIHKFRD 379
 DB 359 CPHNPASFRFLYSLENDIHKFRD 381
 359 CPHNPASFRFLYSLENDIHKFRD 381

RESULT 4
 CE10_CHICK STANDARD; PRT; 375 AA.
 AC P19336;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE CEF-10 protein precursor.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OK NCBI_TaxID=9031;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89145206; PubMed=2537491;
 RA Simmons D.L., Levy D.B., Yannoni Y., Erikson R.L.;
 RT Identification of a phorbol ester-repressible v-src-inducible gene.;
 RL Proc. Natl. Acad. Sci. U.S.A. 86:1178-1182(1989).
 CC -1- FUNCTION: PROBABLE SECRETED REGULATORY PROTEIN.
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- INDUCTION: By v-src.
 CC -1- SIMILARITY: Contains 1 IGF1R domain.
 CC -1- SIMILARITY: Contains 1 VWF domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: J04496; AAA48661.1; -
 DR PIR: A41428; A41428.
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot.C.
 DR InterPro: IPR000867; Ins1_gro_fac-pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; VWF.C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00219; IGF1R; 1.
 DR Pfam: PF00090; TSP_1; 1.
 DR Pfam: PF00093; VWC; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF_BINDING; 1.
 DR PROSITE; PS50092; TSP1; 1.
 DR PROSITE; PS01208; VWF; 1; 1.
 DR PROSITE; PS50184; VWF_2; 1.
 KM Growth factor binding; Signal.
 FT SIGNAL; 1 22
 FT CHAIN; 23 375
 FT DOMAIN; 98 164
 FT DOMAIN; 223 268
 FT DOMAIN; 281 318
 FT DISULFID; 281 318
 FT DISULFID; 298 332
 FT DISULFID; 309 348
 FT DISULFID; 312 350
 FT DISULFID; 317 354
 SQ SEQUENCE 375 AA; 40651 MW; 95F28533BE35D5AE CRC64;

Query Match 78.3%; Score 1646.5; DB 1; Length 375;
 Best Local Similarity 80.1%; Pred. No. 2.3e-114;
 Matches 302; Conservative 22; Mismatches 36; Indels 17; Gaps 7;

QY 10 AVALVTLHLTRLALSTCPACCHCPLEAPKCAPGVGLVDDGCGCCVCAKOLNECCSKQ 68
 10 AVALVTLHLTRLALSTCPACCHCPLEAPKCAPGVGLVDDGCGCCVCAKOLNECCSKQ 68
 DB 9 ALAVALVTLHLTRLALSTCPACCHCPLEAPKCAPGVGLVDDGCGCCVCAKOLNECCSKQ 68
 9 ALAVALVTLHLTRLALSTCPACCHCPLEAPKCAPGVGLVDDGCGCCVCAKOLNECCSKQ 68
 QY 69 PCDHRTKGLKCNFGASSTALKGICRAQSGRCCEVNSRTIYONGESQPNCKHOCTCIDGAV 128
 69 PCDHRTKGLKCNFGASSTALKGICRAQSGRCCEVNSRTIYONGESQPNCKHOCTCIDGAV 128
 DB 69 PCDHRTKGLKCNFGASSTALKGICRAQSGRCCEVNSRTIYONGESQPNCKHOCTCIDGAV 128
 69 PCDHRTKGLKCNFGASSTALKGICRAQSGRCCEVNSRTIYONGESQPNCKHOCTCIDGAV 128
 QY 129 GCIPCLPOELSLPNLGNCPNRLVKVSGGCEWVDEDSIKDSLDDODL--IGLDAS 184
 129 GCIPCLPOELSLPNLGNCPNRLVKVSGGCEWVDEDSIKDSLDDODL--IGLDAS 184
 DB 129 GCIPCLPOELSLPNLGNCPNRLVKVSGGCEWVDEDSIKDSLDDODL--IGLDAS 186
 129 GCIPCLPOELSLPNLGNCPNRLVKVSGGCEWVDEDSIKDSLDDODL--IGLDAS 186
 QY 185 EVELTRNNELIAGKSSILKRLPVFGTEPRVLFNPLAHAGOKCIYQVTSWQSKSCG 242
 185 EVELTRNNELIAGKSSILKRLPVFGTEPRVLFNPLAHAGOKCIYQVTSWQSKSCG 242

```

Db 187 EGEILRNELLAIVKG-GIKMLPVFGSEPOSRAPENP-----KCIYQTTSMSCSKTCG 239
QY 243 TGISTRTYNDNPECLVKEPRICPCGQPVYSSILKGGKCKSTKKSPEVRYTAVGCS 302
Db 240 TGISTRTYNDNPECLVKEPRICPCGQPVYSSILKGGKCKSTKKSPEVRYTAVGCS 299
QY 303 SVKRYRKYCGSCVDGRCCTPLQRTYKMRFCDEGEMFSKNVMIOSCKNCYCPHNE 362
Db 300 SVKRYRKYCGSCVDGRCCTPLQRTYKMRFCDEGEMFSKNVMIOSCKNCYCPHNE 359
QY 363 ASFRYLSLFDIHKFRD 379
Db 360 A-YFRLVNDIHKFRD 375

RESULT 5
CTGF_RAT STANDARD: PRT; 347 AA.
ID CTGF_RAT STANDARD: PRT; 347 AA.
AC Q9RLE9: O9WVSI;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Connective tissue growth factor precursor.
GN CTGF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20145935; PubMed=10679821;
RA Xu J., Smock S.L., Safadi F.F., Rosenzweig A.B., Odgren P.R.,
RT Marks S.C. Jr., Owen T.A., Popoff S.N.:
RT "Cloning the full-length cDNA for rat connective tissue growth factor:
RT implications for skeletal development."
RL J. Cell. Biochem. 77:103-115(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Tezuka K., Tamatani T.:
RT "Rattus norvegicus connective tissue growth factor."
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATRACTANT SECRETED BY
CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL
CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
CC SYNTHESIS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF120275; AAA39132.1; -
DR EMBL: AB023068; BAB82125.1; -
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot_C.
DR InterPro: IPR000867; Insl_gro_fac_pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF000219; IGFBP; 1.
DR Pfam: PF00090; TSP_1; 1.
DR Pfam: PF00093; VWF; 1.
DR SMART: SM00041; CT; 1.

```

```

DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; VWF; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00222; IGFBP; 1.
DR PROSITE: PS50092; TSP1; 1.
DR PROSITE: PS01208; VWF_C; 1.
DR PROSITE: PS50184; VWF_C_2; 1.
DR Cell adhesion; DNA synthesis; Extracellular matrix; Signal.
KM SIGNAL 1 24
FT CHAIN 25 347
FT DOMAIN 31 98
FT DOMAIN 99 165
FT DOMAIN 196 241
FT DOMAIN 254 328
FT DISULFID 254 291
FT DISULFID 271 305
FT DISULFID 282 321
FT DISULFID 285 323
FT DISULFID 290 327
FT DISULFID 35 35
FT CONFLICT 94 94
FT CONFLICT 94 94
SQ SEQUENCE 347 AA; 37756 MW; CFEI1A197687B16 CRC64;

Query Match. 45.6%; Score 960; DB 1; Length 347;
Best local similarity 46.4%; Pred. No. 7.5e-64;
Matches 176; Conservative 61; Mismatches 104; Indels 38; Gaps 7;

QY 1 MSSSTFRLAVAVTLHLRLAL-STCPAACHCLE-APKCAPGVGLVRCGCGCKVCAR 58
Db 1 MLASVAPVSLVALVLLCTRPATGQDCSAQCAEAAPRCBPAGVSLVLDGCGCCRVAC 60
QY 59 QLNDDCSKTQPCDHTKGLGECFEGASSSTALKICIAQOSGRCEFNRYONGESFQPNCK 118
Db 61 QLGELCTERPCDCHKGLFCDGSPANKRIGVCTAK-DGACVGVGSYRGSSEFSQSC 119
QY 119 HOCICIGAVGCIPLCPQELSLPMLGCPNPLVYVSGCCCEWCEDESIDDDODL 178
Db 120 YQCTCLDGAAGCVPLCSMDVRLPSPDCFPFRVRLPGKCEWVCDPE----- 167
QY 179 LGLDASEVELRNELLAIGKSSIKRL-PVFGTEPRVLPNLAHAGKCIYQTTSMSC 237
Db 168 -----KDRIVYGALAAVRLIEDFGDPIMM-----RANCLVQTTSMSC 207
QY 238 SKSCGTGISTRTYNDNPECLVKEPRICPCGQPVYSSILKGGKCKSTKKSPEVRYT 297
Db 208 SKTCGMGISTRTYNDNPECLVKEPRICPCGQPVYSSILKGGKCKSTKKSPEVRYT 267
QY 298 YAGGSVKKYRKYRKYCGSCVDGRCCTPLQRTYKMRFCDEGEMFSKNVMIOSCKNCY 357
Db 268 LSGCTSVKTYRKYRKYCGSCVDGRCCTPLQRTYKMRFCDEGEMFSKNVMIOSCKNCY 327
QY 358 PHPNEASFRLY--SLFNDI 374
Db 328 PGDMIDRESILYRKMGDM 346

RESULT 6
CTGF_MOUSE STANDARD: PRT; 348 AA.
ID CTGF_MOUSE STANDARD: PRT; 348 AA.
AC P29268: 0922U0;
DT 28-FEB-1992 (Rel. 24, Created)
DT 15-SEP-2003 (Rel. 41, Last sequence update)
DE Connective tissue growth factor precursor (FISP-12 protein)
DE (Hypertrophic chondrocyte-specific protein 24).
GN CTGF OR FISP12 OR FISP-12 OR HCS24.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

RP SEQUENCE FROM N.A.
 RA MEDLINE-91363290; PubMed-1888698;
 RA Rybeck R.-P., Macdonald-Bravo H., Mattei M.-G., Bravo R.;
 RT "Structure, mapping, and expression of fisp-12, a growth factor-
 RT inducible gene encoding a secreted cysteine-rich protein.";
 RL Cell Growth Differ. 2:225-233(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-91229699; PubMed-2029337;
 RA Brunner A., Chinn J., Neubauer M.G., Purchio A.F.;
 RT "Identification of a gene family regulated by transforming growth
 RT factor-beta.";
 RL DNA Cell Biol. 10:293-300(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-22388257; PubMed-12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stalton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Baha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fehey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmitt J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP FUNCTION AND SUBCELLULAR LOCATION.
 RA MEDLINE-97327410; PubMed-9184077;
 RA Kireeva M.L., Iatinkic B.V., Kolesnikova T.V., Chen C.C., Yang G.P.,
 RA Adler A.S., Lau L.F.;
 RT "Cyr61 and fisp2 are both ECM-associated signaling molecules:
 RT activities, metabolism, and localization during development.";
 RL Exp. Cell Res. 233:63-77(1997).
 RN [5]
 RP FUNCTION.
 RA MEDLINE-99182484; PubMed-10082563;
 RA Babic A.M., Chen C.C., Lau L.F.;
 RT "Fisp2/mouse connective tissue growth factor mediates endothelial
 RT cell adhesion and migration through integrin alphavbeta3, promotes
 RT endothelial cell survival, and induces angiogenesis in vivo.";
 RL Mol. Cell. Biol. 19:2958-2966(1999).
 CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATRACTANT SECRETED BY
 CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
 CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL
 CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
 CC SYNTHESIS (BY SIMILARITY).
 CC -1- SUBUNIT: Monomer (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
 CC SOLUBLE FORM.
 CC -1- TISSUE SPECIFICITY: TESTIS, SPLEEN, KIDNEY, LUNG, HEART, AND BRAIN
 CC (LOWEST LEVEL IN TESTIS AND HIGHEST IN LUNG).
 CC -1- INDUCTION: By growth factors.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 WFC domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M70641; AAA37627.1; -;
 CC DR EMBL: M70642; AAA37628.1; -;
 CC DR EMBL: M80263; AAA37135.1; -;
 CC DR EMBL: BC006783; AA06783.1; -;
 CC DR PIR: A40578; AA0578.
 CC DR MGI: 95537; Ctgf.
 CC DR GO: 0005578; C:extracellular matrix; IDA.
 CC DR GO: 0008201; F:heparin binding activity; IDA.
 CC DR GO: 0005178; F:integrin binding activity; IDA.
 CC DR GO: 0001525; P:angiogenesis; IDA.
 CC DR GO: 0001647; P:cell migration; IDA.
 CC DR GO: 0007160; P:cell-matrix adhesion; IDA.
 CC DR GO: 0008543; P:FGF receptor signaling pathway; IDA.
 CC DR GO: 0007229; P:integrin-mediated signaling pathway; IDA.
 CC DR GO: 0001503; P:proliferation; IMP.
 CC DR InterPro: IPR006208; Cys_knot.
 CC DR InterPro: IPR006207; Cys_knot_C.
 CC DR InterPro: IPR000867; Insl_gro_fac-dr.
 CC DR InterPro: IPR000864; TSP1.
 CC DR InterPro: IPR001007; WFC_C.
 CC DR Pfam: PF00007; Cys_knot; 1.
 CC DR Pfam: PF00219; IGFBP; 1.
 CC DR Pfam: PF00090; tsp.1; 1.
 CC DR Pfam: PF00093; wvc; 1.
 CC DR SMART: SM00041; CT; 1.
 CC DR SMART: SM00121; IB; 1.
 CC DR SMART: SM00209; TSP1; 1.
 CC DR SMART: SM00214; WVC; 1.
 CC DR PROSITE: PS01185; CTCK_1; 1.
 CC DR PROSITE: PS01225; CTCK_2; 1.
 CC DR PROSITE: PS00222; TGF_BINDING; 1.
 CC DR PROSITE: PS00082; TSP1; 1.
 CC DR PROSITE: PS01208; WFC_1; 1.
 CC DR PROSITE: PS0184; WFC_2; 1.
 CC DR Cell adhesion; DNA synthesis;
 CC FT SIGNAL; 1 25
 CC FT CHAIN; 26 348
 CC FT DOMAIN; 32 99
 CC FT DOMAIN; 100 166
 CC FT DOMAIN; 197 242
 CC FT DOMAIN; 255 329
 CC FT DISULFID; 255 292
 CC FT DISULFID; 272 306
 CC FT DISULFID; 283 322
 CC FT DISULFID; 286 324
 CC FT DISULFID; 291 328
 CC FT CONFLICT; 161 161
 CC SQ SEQUENCE 348 AA; 37794 MW; 4D7B6D9089174049 CRC64;
 CC
 CC Query Match 45.5%; Score 957.5; DB 1; Length 348;
 CC Best Local Similarity 46.8%; Pred. No. 1,1e-63;
 CC Matches 178; Conservative 60; Mismatches 103; Indels 39; Gaps 8;
 CC
 CC 1 MESSRTTAAVAVTILHL-TRLAL-STCPAACHPLE-APKCAPGAYVRDCCGCKCYA 57
 CC 1 MASVAGPISLVLALCTRPATGDCSAQCQCAEAAPHCAPAGSLVLDGCGCRVCA 60
 CC
 CC 58 KQINEDCKTOPCDHTKGLGECNFGASSTALKICRQSGRCEYNSRYONGSEFOPNC 117
 CC 61 KQLGELCTRPDCDPKHLGFCDFGSPANKKIGVCTAK-DGACVCGGSGYRGESEFQSSC 119
 CC
 CC 118 KQCTCIDGAVGCIPLQPELSPLNGCPNPLVVGQCCGEWVCDEDSIKDSLDDDD 177
 CC 120 KYQCTCLDGAVGCVPLCSMDVRLPSPDCPFPRVRLPKCCGEWVCDEP----- 168
 CC
 CC 178 LIGLDASVEELRNNEMLAIGKSSLKRLL-PVEGTEPRVLPNPLAHQOKIVOTTSNQ 236
 CC 169 -----KRTAVGPAALAAVRLDTEGDPDTM-----RANCLVOTTEWSA 207
 CC 237 CSKSGGTGISTRTVNDNPECRILVKETRICEVAPCGQPYVSSILKKGKSKTKKSPPEVR 296


```

Db 208 CSKTCGAGISTRYNDWTFCELEKOSRLCMWRPCPCADLEENIKKGGKCC:RTPKIAKPYKE 267
Qy 297 TYACSSVKKYRPRYCCSCVDGRCCTPLQRTYVARRRCEDEGEFSSNNYMIQCKKNYN 356
Db 268 ELSCGTSVKYKRAFCVCTDGRCCPTHRRTTLFVEFKCPDGEIMKKNMEIKYCACHYN 327
Qy 357 CPHNEASFRLY--SLFNDI 374
Db 328 CPQNDLFESELYIKRMKGDM 347

RESULT 7
CTGF_HUMAN STANDARD; PRT; 349 AA.
AC P29279; Q960X2;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Connective tissue growth factor precursor (Hypertrophic chondrocyte-
DE specific protein 24).
CN CTGF OR HCS24.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Umbilical vein endothelial cells;
RX MEDLINE=91373462; PubMed=1654338;
RA Bradham D.M., Igarashi A., Potter R.L., Grotendorst G.R.;
RT "Connective tissue growth factor: a cyclin-dependent kinase secreted by
RT human vascular endothelial cells is related to the SRC-induced
RT immediate early gene product CEF-10."
RL J. Cell Biol. 114:1285-1294(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Umbilical vein endothelial cells;
RX MEDLINE=93187114; PubMed=1293144;
RA Igarashi A., Bradham D.M., Okochi H., Grotendorst G.R.;
RT "Connective tissue growth factor."
RL J. Dermatol. 19:642-643(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97207446; PubMed=9054739;
RA Oemar B.S., Werner A., Garnier J.M., Do D.D., Godoy N., Nuuck M.,
RA Matz W., Rupp J., Pech M., Luescher T.F.;
RT "Human connective tissue growth factor is expressed in advanced
RT atherosclerotic lesions."
RL Circulation 95:831-839(1997).
RN [4]
RP SEQUENCE FROM N.A.
RA COOLEY V.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP FUNCTION.
RC TISSUE-Chondrocytes;
RX MEDLINE=20080284; PubMed=10614647;
RA Nakanishi T., Nishida T., Shimo T., Kobayashi K., Kubo T.,
RA Tamatani T., Tezuka K., Takigawa M.;
RT "Effects of CTGF/HCS24, a product of a hypertrophic chondrocyte-
RT specific gene, on the proliferation and differentiation of
RT chondrocytes in culture."
RL Endocrinology 141:264-273(2000).
RN [6]
RP FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTPRACANT SECRETED BY
RN [7]
RP VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
RN [8]
RP DIFFERENTIATION OF CHONDROCYTES. MEDIATES CELL ADHESION AND
RN [9]
RP ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA SYNTHESIS.
RN [10]
RP SUBUNIT: Monomer.
RN [11]
RP SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
RN [12]
RP SOLUBLE FORM (BY SIMILARITY).
RN [13]
RP ALTERNATIVE PRODUCTS:
RN [14]
RP Event-Alternative splicing; Named isoforms=2;

```

```

CC CC Name=Long;
CC CC IsoId=P29279-1; Sequence=Displayed;
CC CC Name=Short;
CC CC IsoId=P29279-2; Sequence=VSP_002460;
CC CC Note-No experimental confirmation available;
CC CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC CC -1- SIMILARITY: Contains 1 WFBC domain.
CC CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC CC EMBL; M92934; AAA91279.1;
CC CC EMBL; X78947; CAA5544.1;
CC CC EMBL; A1354866; CAC44023.1;
CC CC PIR; A40551; A40551.
CC CC Genew; HGNC:2500; CTGF.
CC CC MIM; 121009;
CC CC GO; GO:0005578; C:extracellular matrix; TAS.
CC CC GO; GO:0005886; C:plasma membrane; TAS.
CC CC GO; GO:0005520; F:insulin-like growth factor binding activity; TAS.
CC CC GO; GO:0008151; P:cell growth and/or maintenance; TAS.
CC CC GO; GO:0008544; P:epidermal differentiation; TAS.
CC CC GO; GO:0009611; P:response to wounding; TAS.
CC CC InterPro; IPR006208; Cys_knot.
CC CC InterPro; IPR006207; Cys_knot_C.
CC CC InterPro; IPR000867; Insl_gro_fac_pr.
CC CC InterPro; IPR000884; TSP1.
CC CC Pfam; PF00007; Cys_knot; 1.
CC CC Pfam; PF00219; IGFBP; 1.
CC CC Pfam; PF00090; tsp_1; 1.
CC CC Pfam; PF00093; wvc; 1.
CC CC SMART; SM00041; CT; 1.
CC CC SMART; SM00121; IB; 1.
CC CC SMART; SM00209; TSP1; 1.
CC CC SMART; SM00214; WVC; 1.
CC CC PROSITE; PS01185; CTCK_1; 1.
CC CC PROSITE; PS01225; CTCK_2; 1.
CC CC PROSITE; PS00222; IGF_BINDING; 1.
CC CC PROSITE; PS00092; TSP1; 1.
CC CC PROSITE; PS01208; WFBC_1; 1.
CC CC PROSITE; PS01084; WFBC_2; 1.
CC CC Cell adhesion; DNA synthesis; Extracellular matrix; Signal;
CC KW Alternative splicing.
CC FT SIGNAL 1 26
CC FT CHAIN 27 349
CC FT DOMAIN 33 100
CC FT DOMAIN 101 167
CC FT DOMAIN 198 243
CC FT DOMAIN 256 330
CC FT DISULFID 256 293
CC FT DISULFID 273 307
CC FT DISULFID 284 323
CC FT DISULFID 287 325
CC FT DISULFID 292 329
CC FT CARBOHYD 28 28
CC FT CARBOHYD 225 225
CC FT VARSPLIC 172 198
CC FT CONFLICT 83 83
CC FT SEQUENCE 349 AA; 38069 MW; 0ECF8470B357EA95 CRC64;
CC Query Match 45.28; Score 950; DB 1; Length 349;
CC Best Local Similarity 45.78; Pred. No. 4; le-63;
CC Matches 176; Conservative 64; Mismatches 97; Indels 48; Gaps 9;

```

```

QY 1 MSSSTFRTLAVALTLHLTRIALSTCPAA-----CHCPLE-APKCAPGVGLVRDCCG 52
DB 1 MTAASMPRVAAFEVYL-----LALCSRPVAGONCSGPCRCDPEAPRCAPAGSLVLDGCG 56
QY 53 CRYCAKOLNEDSCSTOPOCDHTKGLCECNFGASSALKICCAOSEGRCEYNSRIYONGES 112
DB 57 CRYCAKOLNEDSCSTOPOCDHTKGLCECNFGASSALKICCAOSEGRCEYNSRIYONGES 115
QY 113 FQPNCKHOCTCIDGAVGICPLCPQELSLPNLGCNPRLVYVSGOCCEWVCDSDISL 172
DB 116 FQSSCKYQCTCIDGAVGICPLCPQELSLPNLGCNPRLVYVSGOCCEWVCDSDISL 170
QY 173 DDQDGLIGDASEVELETRNNELIAGKSSILKRL-PVETGEPRVLEPNLHAHQKCIYOT 231
DB 171 -DQ-----TIVGPAALAYRLEDETFGDPPTMI-----RANCLVOT 203
QY 232 TSMSCSKSGTGISTRTVNDNPECRLYKTRICEVRPGQPYVSSLKKGKSKTKKSP 291
DB 204 TSMSCSKSGTGISTRTVNDNPECRLYKTRICEVRPGQPYVSSLKKGKSKTKKSP 293
QY 292 EPVRFYAGSSVKKYRPKYCGSCVDGRCTPLQTRTKKFRCEDEGEMFSKNVMYIOSC 351
DB 264 KPIKPELSSGCTSKMTYAKKCGVCTDGRCTPHRTTLVPEFRCPDGEVKKMMFTKIC 323
QY 352 KCNYNCPHNEASFRIL--SLFNDI 374
DB 324 ACHYNCPCGNDIFESLYRKMGM 348

```

RESULT 8

```

CTGF_BOVIN STANDARD; PRT; 349 AA.
ID CTGF_BOVIN STANDARD; PRT; 349 AA.

```

```

AC 018739; 09GL71;
DT 15-JUL-1998 (Rel. 36, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Connective tissue growth factor precursor.
GN CTGF.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Arteria;
RA Lillensiek B., Lin Z., Fotsis T., Schimanski M., Bierhaus A.,
RA Kanitz M., Kauffmann G., Schweigerer L., Ziegler R., Nawroth P.P.;
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Mathias M., Schwitters C., Hove M., Rupp S., Erondu N.E.;
RT "Bovine connective tissue growth factor, organization of the
RT chromosomal gene and demonstration of promoter activity.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATRACTANT SECRETED BY
CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL
CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
CC SYNTHESIS (BY SIMILARITY).
CC -1- SUBUNIT: Monomer (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
CC SOLUBLE FORM (BY SIMILARITY).
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WFC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib.ch).
CC -----
DB 1 MSSSTFRTLAVALTLHLTRIALST-----CPAACHPL-BAKCAPGVGLVRDCCG 52
DB 1 MSATGIGPVRCAFEVYL-----LALCSRPVAGONCSGPCRCDPEAPRCAPAGSLVLDGCG 56
QY 53 CRYCAKOLNEDSCSTOPOCDHTKGLCECNFGASSALKICCAOSEGRCEYNSRIYONGES 112
DB 57 CRYCAKOLNEDSCSTOPOCDHTKGLCECNFGASSALKICCAOSEGRCEYNSRIYONGES 115
QY 113 FQPNCKHOCTCIDGAVGICPLCPQELSLPNLGCNPRLVYVSGOCCEWVCDSDISL 172
DB 116 FQSSCKYQCTCIDGAVGICPLCPQELSLPNLGCNPRLVYVSGOCCEWVCDSDISL 169
QY 173 DDQDGLIGDASEVELETRNNELIAGKSSILKRL-PVETGEPRVLEPNLHAHQKCIYOT 231
DB 170 -----KEHTVGPALAAVRPEDETFGDPPTMI-----RANCLVOT 203
QY 232 TSMSCSKSGTGISTRTVNDNPECRLYKTRICEVRPGQPYVSSLKKGKSKTKKSP 291
DB 204 TSMSCSKSGTGISTRTVNDNPECRLYKTRICEVRPGQPYVSSLKKGKSKTKKSP 293
QY 292 EPVRFYAGSSVKKYRPKYCGSCVDGRCTPLQTRTKKFRCEDEGEMFSKNVMYIOSC 351
DB 264 KPIKPELSSGCTSKMTYAKKCGVCTDGRCTPHRTTLVPEFRCPDGEVKKMMFTKIC 323
QY 352 KCNYNCPHNEASFRIL--SLFNDI 374
DB 324 ACHYNCPCGNDIFESLYRKMGM 348

```

Query Match 44.6%; Score 937; DB 1; Length 349;

Best local Similarity 44.7%; Pred. No. 3.7e-62;

Matches 172; Conservative 65; Mismatches 100; Indels 48; Gaps 8;

```

QY 1 MSSSTFRTLAVALTLHLTRIALST-----CPAACHPL-BAKCAPGVGLVRDCCG 52
DB 1 MSATGIGPVRCAFEVYL-----LALCSRPVAGONCSGPCRCDPEAPRCAPAGSLVLDGCG 56
QY 53 CRYCAKOLNEDSCSTOPOCDHTKGLCECNFGASSALKICCAOSEGRCEYNSRIYONGES 112
DB 57 CRYCAKOLNEDSCSTOPOCDHTKGLCECNFGASSALKICCAOSEGRCEYNSRIYONGES 115
QY 113 FQPNCKHOCTCIDGAVGICPLCPQELSLPNLGCNPRLVYVSGOCCEWVCDSDISL 172
DB 116 FQSSCKYQCTCIDGAVGICPLCPQELSLPNLGCNPRLVYVSGOCCEWVCDSDISL 169
QY 173 DDQDGLIGDASEVELETRNNELIAGKSSILKRL-PVETGEPRVLEPNLHAHQKCIYOT 231
DB 170 -----KEHTVGPALAAVRPEDETFGDPPTMI-----RANCLVOT 203
QY 232 TSMSCSKSGTGISTRTVNDNPECRLYKTRICEVRPGQPYVSSLKKGKSKTKKSP 291
DB 204 TSMSCSKSGTGISTRTVNDNPECRLYKTRICEVRPGQPYVSSLKKGKSKTKKSP 293
QY 292 EPVRFYAGSSVKKYRPKYCGSCVDGRCTPLQTRTKKFRCEDEGEMFSKNVMYIOSC 351
DB 264 KPIKPELSSGCTSKMTYAKKCGVCTDGRCTPHRTTLVPEFRCPDGEVKKMMFTKIC 323
QY 352 KCNYNCPHNEASFRIL--SLFNDI 374
DB 324 ACHYNCPCGNDIFESLYRKMGM 348

```

```

Db 264 KPFFELSGCTSMYTAKEGVCCTDRCCPHRTTLLPVEFKCPDEP/MAKSMF1KTC 323
QY 352 KCNTNCPHPNEASFRLY--SLFNDI 374
Db 324 ACHYNCPGDNDIFESLYRKMYGDM 348

RESULT 9
CTGF_PIG STANDARD; PRT; 349 AA.
ID CTGF_PIG
AC 019113;
DR 15-JUL-1998 (Rel. 36, Created)
DR 15-JUL-1998 (Rel. 36, Last sequence update)
DR 28-FEB-2003 (Rel. 41, Last annotation update)
DE Connective tissue growth factor precursor.
GN CTGF.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9623;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Uterus;
RX MEDLINE=97390475; PubMed=9242708;
RA Briggs D.R., Steffen C.L., Kim G.Y., Vegunta R.K., Diehl J.R.,
RA Harding P.A.;
RT "Purification and characterization of novel heparin-binding growth
RT factors in uterine secretory fluids. Identification as heparin-
RT regulated Mr 10,000 forms of connective tissue growth factor.";
RL J. Biol. Chem. 272:20275-20282(1997).
CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL
CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
CC SYNTHESIS (BY SIMILARITY).
CC -1- SUBUNIT: Monomer (BY similarity).
CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
CC SOLUBLE FORM (BY SIMILARITY).
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WPC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1st-sib.ch/announce/
CC or send an email to license@1st-sib.ch).
CC
DR EMBL, U83916; AAC48756.1; -
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot_C.
DR InterPro: IPR000867; Insl_gro_fac_pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; WVF_C.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; tsp_1; 1.
DR Pfam: PF00093; WVC; 1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; WVC; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS00093; TSP1; 1.
DR PROSITE: PS01208; WVF_C; 1.
DR PROSITE: PS0184; WVF_C_2; 1.
KW Cell adhesion; DNA synthesis; Extracellular matrix; Signal.

```

```

FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 349 CONNECTIVE TISSUE GROWTH FACTOR.
FT DOMAIN 33 100 IGFBP.
FT DOMAIN 101 167 WVC.
FT DOMAIN 198 243 TSP TYPE-1.
FT DOMAIN 256 330 CTCK.
FT DISULFID 256 293 BY SIMILARITY.
FT DISULFID 273 307 BY SIMILARITY.
FT DISULFID 284 323 BY SIMILARITY.
FT DISULFID 287 325 BY SIMILARITY.
FT DISULFID 292 329 BY SIMILARITY.
SO SEQUENCE 349 AA; 38007 MW; BB510E2B2B52D4A0 CRC64;

Query Match 43.7% Score 919; DB 1; Length 349;
Best Local Similarity 44.9%; Pred. No. 7.9e-61;
Matches 173; Conservative 62; Mismatches 102; Indels 48; Gaps 9;

QY 1 MSSSTFRTLLAVAVLLHLTRLALSTCPAA-----CHCPL-EAPKCAPGVGLVRDCCGC 52
Db 1 MSATGLSPVRCAPVLL-----LALCSRPASGGDCGCGCAAGKRACRACGATVLDGCGC 56
QY 53 CKVCAKQDNEDCKTQPCDHTKGLKCNFGASSTALKGICRARSRGRCETNSRYQNGES 112
Db 57 CRLCAKQDGLGELCTERDCDHPKGLFCDFGSPANKKIGVCTAK-DGAPCVGEGTYRRSGES 115
QY 113 FQPNCKHOCICIDAVGCIPLCPQELSLPNIKGNPNLYKVGSGCCGEMWCDDESINDSL 172
Db 116 FQSSCKYQCTCLDGAAGCVPLCSMDVRLPSPDCPPFRVVKLPGKCEWVWDEP--KDH- 172
QY 173 DDQDGLGLDASEVELTRNNELIAGKSSLRKL-PVGEPEPRVLPFLNHAHGOKCIYQT 231
Db 173 -----TYVGPALAAVRLIEDTFGPDPTMM-----RANCLVQT 203
QY 232 TSMSCSKSGCTGISTRYTNDNPECLVKEIRICEVRPGQPVYSSLKKGKSKYTKSP 291
Db 204 TEMSACSKTCGCMGISTRYTNDNAPCRLEKOSRLCMVBPCEADLEENIKKKKIRTPKIS 263
QY 292 EPRVETIAGGSSVYKYPKYGCGSCVDRCPPLOTFTVKKMFRCDEDEMSKNYMMIOSC 351
Db 264 KPFFELSGCTSMYTAKEGVCCTDRCCPHRTTLLPVEFKCPDEP/MAKSMF1KTC 323
QY 352 KCNTNCPHPNEASFRLY--SLFNDI 374
Db 324 ACHYNCPGDNDIFESLYRKMYGDM 348

RESULT 10
NOV_CHICK STANDARD; PRT; 351 AA.
ID NOV_CHICK
AC P28686;
DR 01-DEC-1992 (Rel. 24, Created)
DR 01-DEC-1992 (Rel. 24, Last sequence update)
DR 28-FEB-2003 (Rel. 41, Last annotation update)
DE NOV protein precursor (Nephroblastoma overexpressed gene protein).
GN NOV.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Brown Leghorn;
RX MEDLINE=92107157; PubMed=1309586;
RA Joliet V., Marlierie C., Dambirine G., Plaastert G., Brissac M.,
RA Crochet J., Perbal B.;
RT "Proximal rearrangements and overexpression of a new cellular gene
RT (nov) in myeloblastosis-associated virus type 1-induced
RT nephroblastomas";
RL Mol. Cell. Biol. 12:10-21(1992).
CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
CC GROWTH REGULATION. ITS OVEREXPRESSION IS ASSOCIATED WITH
CC TUMORIGENESIS AND EXPRESSION OF A N-TERMINAL-TRUNCATED VERSION

```

```

CC OF NOV GENE IN CHICKEN EMBRYONIC FIBROBLASTS (CEF) IS SUFFICIENT
CC TO INDUCE THE TRANSFORMATION OF CEF IN VITRO.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: BRAIN AND HEART, AND AT A LOWER LEVEL IN
CC MUSCLE AND INTESTINE, IN THE EMBRIO. LONG AND LESS SO IN BRAIN AND
CC SPLEEN, IN ADULT CHICKEN.
CC -1- DEVELOPMENTAL STAGE: MAV1-INDUCED NEPHROBLASTOMAS EXPRESS A HIGH
CC LEVEL OF NOV GENE WHOSE TRANSCRIPTION IS NORMALLY ARRESTED IN
CC ADULT KIDNEY.
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WFRC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X59284; CAA1975.1; -
DR PIR: S20078; S20078.
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot_C.
DR InterPro: IPR000867; Insl_gro_fac_pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; WMF_C.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; TSP_1; 1.
DR Pfam: PF00093; WWC; 1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; WWC; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS50092; TSP1; 1.
DR PROSITE: PS01208; WMF_1; 1.
DR PROSITE: PS50184; WMF_2; 1.
DR PROTO-ONCOGENE; Growth factor; Signal.
KW SIGNAL.
FT CHAIN 1 24 POTENTIAL.
FT DOMAIN 25 351 NOV PROTEIN.
FT DOMAIN 31 103 IGFBP.
FT DOMAIN 104 170 WMF.
FT DOMAIN 201 246 TSP TYPE-1.
FT DOMAIN 258 332 CTCK.
FT DISULFID 258 295 BY SIMILARITY.
FT DISULFID 275 309 BY SIMILARITY.
FT DISULFID 286 325 BY SIMILARITY.
FT DISULFID 289 327 BY SIMILARITY.
FT DISULFID 294 331 BY SIMILARITY.
FT CARBOHYD 274 274 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 351 AA; 38268 MW; 1ECB3FA3058C6797 CRC64;
Query Match 40.9%; Score 861; DB 1; Length 351;
Best Local Similarity 44.5%; Pred. No. 1.5e-56;
Matches 165; Conservative 44; Mismatches 110; Indels 52; Gaps 7;

```

```

QY 176 DLLGLDASEVELTRNNELIAIGKSLRLPVFTEPRVLEPNPLAHGOKCIYVOTTS 235
   1 : : : : :
Db 188 EATLGLDVSD-----SSANCEIQTETS 210
QY 236 QCSKSCGTISTRTVNDNPECLVKEFRICENRPGCPYSSLKKGKCKSKTKKSPYR 295
   1 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 211 ACSKSCGMSFTVTRNRNOOCEWVKOTRLCMRPPCENEPSD-KKGKCIQTIKKSKAVR 269
QY 296 FTYAGSSVKKRPRKCGSCVNGRCCTPLOTFRVKRFRCEGCEMSKRVMMIOGCKAV 355
   1 : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 270 FEYKNTSVQTKRPRYGLCNDGRCTPHNTKTIQVERCPDGKFLKPMMLINTVCVG 329
QY 356 NCPHPNEASFR 366
   1 : : : : :
Db 330 NCPQSNNAFFQ 340.

RESULT 11
NOV_COTJA STANDARD; PRT; 353 AA.
AC P42642:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
GN NOV protein precursor (Nephroblastoma overexpressed gene protein).
OS coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_Taxid=93934;
RN [1]
RP SEQUENCE FROM N.A.
RA Weiskirchen R., Bister K.;
RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
CC GROWTH REGULATION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WFRC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U13063; AAA2128.1; -
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot_C.
DR InterPro: IPR000867; Insl_gro_fac_pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; WMF_C.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; TSP_1; 1.
DR Pfam: PF00093; WWC; 1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; WWC; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS50092; TSP1; 1.
DR PROSITE: PS01208; WMF_1; 1.
DR PROSITE: PS50184; WMF_2; 1.
KW PROTO-ONCOGENE; Growth factor; Signal.

```


DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF_BINDING; 1.
 DR PROSITE; PS50092; TSP1; 1.
 DR PROSITE; PS01208; VWC_1; 1.
 DR PROSITE; PS50184; VWC_2; 1.
 KM Growth factor; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 357
 FT DOMAIN 35 107
 FT DOMAIN 108 174
 FT DOMAIN 205 250
 FT DISULFID 264 338
 FT DISULFID 264 301
 FT DISULFID 281 315
 FT DISULFID 292 331
 FT DISULFID 295 333
 FT CARBOHYD 300 337
 FT CARBOHYD 97 97
 FT CARBOHYD 280 280
 FT CONFLICT 97 97
 SQ SEQUENCE 357 AA; 39162 MW; 035D5B4576B85B CRC64;

Query Match 39.9%; Score 839.5; DB 1; Length 357;
 Best Local Similarity 43.8%; Pred. No. 5.8e-55;
 Matches 163; Conservative 52; Mismatches 116; Indels 41; Gaps 9;

QY 1 MSSTF-----RTLAVALTLHLTRALST--CPAAC--HCPLEAPKAPGVGLVRCGCG 52
 DB 4 VQSTFCLRKQKQCLFLFLHLGQVAATQRCPPGRCPPAPPPCAPGVRAVLDCGSC 63
 QY 53 CKVCAROLNDCSKTQPCDHTKGLKCNFGASSFALGICRAQSEGRPCENSRITYNQGES 112
 DB 64 CLVCAQARGESGDLDEPCDESSGLYCDKSDPNSQIGCTA-VEGDNCFYDGYITSGEK 122
 QY 113 FQPNCRHQCTCIDGAVGCIPLCPQELSLPMLGCPNRLKVSQCCCEWVCDSDIKDSL 172
 DB 123 FQPSKFCQCTCRDQIGCVPRCQDVLTPPNCPPARKVEVPECECKMTCGPD-EDSL 181
 QY 173 DDQDGLGLDASVEVLTNNELIAIGKSSKLRLPVGEPRVLFNPLAHGQKCIYQTT 232
 DB 182 GG-----LTLAAYPEATLGEV-----SDSSV-----NCTEQTT 211
 QY 233 SMSGCKSGCTGISTRTVNDNPECRIVKERICEVBRPGQ-PYSSILKKGKSKTKKSP 291
 DB 212 EMTACKSGCMGFSVYTNRNKRCCEMLKQTRLCQVPRCEDEPEQPTDKKCKLRTKSL 271
 QY 292 EPRVFTYAGCSVYKRYPRYCGSCVDGRCCTPLQTRTVKMRFCEDGEMFSKNVMMIOQC 351
 DB 272 KAIHLQFKNCTSLHYTKPRFCGVCSDGRCCTPHNTTIAEFQCSGQIYKKRVMMIQC 331
 QY 352 KCMYNGPHRNEA 363
 DB 332 TCFHNCPPKNEA 343

RESULT 13
 NOV_RAT
 ID NOV_RAT STANDARD; PRT; 351 AA.
 AC 09Q0205;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE NOV protein homolog precursor (Novb) (Nephroblastoma overexpressed gene protein homolog).
 DE gene protein homolog.
 GN NOV.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCBL_TaxID=10116;

RN [1]
 RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
 RC STRAIN-Sprague-Dawley;
 RX MEDLINE=20035752; Pubmed=10570975;
 RA Liu C., Liu X.J., Crowe P.D., Keiner G.S., Fan J., Barry G., Mann F.,
 RA Ling N., De Souza E.B., Maki R.A.;
 RT "Nephroblastoma overexpressed gene (NOV) codes for a growth factor
 RT that induces protein tyrosine phosphorylation."
 RL Gene 238:471-478(1999).
 CC - FUNCTION: Can act as a growth factor for some cells and binds to a
 CC specific receptor that leads to the phosphorylation of a 221 kDa
 CC protein.
 CC - SUBUNIT: Interacts with FBLN1 (By similarity).
 CC - SUBCELLULAR LOCATION: Secreted.
 CC - TISSUE SPECIFICITY: Ubiquitous.
 CC - SIMILARITY: Contains 1 IGF1R domain.
 CC - SIMILARITY: Contains 1 TSP type-1 domain.
 CC - SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: AF171936; AAD49371.1; -
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Ins1_gro_fac_pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; VWC_C.
 DR Pfam: PF000007; Cys_knot; 1.
 DR Pfam: PF000090; IGF1R; 1.
 DR Pfam: PF000090; TSP_1; 1.
 DR Pfam: PF000093; VWC; 1.
 DR SMART; SM00041; CT_1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF_BINDING; 1.
 DR PROSITE; PS50092; TSP1; 1.
 DR PROSITE; PS01208; VWC_1; 1.
 DR PROSITE; PS50184; VWC_2; 1.
 KW Growth factor; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 351
 FT DOMAIN 29 101
 FT DOMAIN 102 168
 FT DOMAIN 199 244
 FT DOMAIN 238 332
 FT DISULFID 238 295
 FT DISULFID 258 295
 FT DISULFID 275 309
 FT DISULFID 286 325
 FT DISULFID 289 327
 FT DISULFID 294 331
 FT CARBOHYD 91 91
 FT CARBOHYD 274 274
 SQ SEQUENCE 351 AA; 38509 MW; 02619707DETCIBFB CRC64;

Query Match 39.4%; Score 829.5; DB 1; Length 351;
 Best Local Similarity 43.9%; Pred. No. 3.1e-54;
 Matches 158; Conservative 51; Mismatches 114; Indels 37; Gaps 8;
 QY 9 LAVALTLHLTRALST--CPAAC--HCPLEAPKAPGVGLVRCGCGCKVCAKOLNDC 64
 DB 10 LCLGFLHLHLNLQVSAITLNCPRCPGSCPSIPTCAPGVRAVLDCSCCPVCAQARGESC 69
 QY 65 SKTPQCDHTKGLKCNFGASSFALGICRAQSEGRPCENSRITYNQGESFQPNCKHQCTCI 124

```

Db 70 SEMRPPDSSGSLVCDRSADPNNGEIGICMV-PEGDNCVFDGVIYRNGEKI-EPNQCQYHCTCR 128
QY 125 DGAVGCTPLCPQELSLPMLGCPNRLKVSAGCCCEEWCEDEDSIKSLDDQDLGLGDAS 184
Db 129 DGOIGCVPRQQLVLLPDPCCPAPKKAAYVGECEKWTGCSSE-KGTLIG----LALPAY 183
QY 185 EVELTRNNELIAGKSSLRKLPVFGTEPRVLEFNPLAHGOKCIYVOTISQSCSKSGCTG 244
Db 184 REPRATVGEV----SDSI------NCTEQTTIMKASCSKSGCMG 217
QY 245 ISTRVNDNDECRVLYETRICEVPRGQ-PVYSLKKKKCKSKTKSPFVRFTYAGCSS 303
Db 218 LSTRVTRNRLQCEWVQKTRICVPRCEQEPGEATDMKGGKCLRTKSKLSIHLQFKNCTS 277
QY 304 VKYIRRYKCGSCVDGRCTPLQTRIVKMRRCEDGEMFSKNVMQISGICNYNCPHPNEA 363
Db 278 LTYTKPRFCIGICSDGRCTPFNRTKIQVEFQCLPQGIKKRPVAVIGTCVCHSNCPQNNEA 337

RESULT 14
NOV_MOUSE STANDARD; PRT; 354 AA.
ID NOV_MOUSE
AC 064239;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE NOV protein homolog precursor (NOVH) (Nephroblastoma overexpressed
  gene protein homolog).
GN NOV.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
  SEQUENCE FROM N.A.
  STRAIN=129/SV, and ICR; TISSUE=Brain;
  MEDLINE=97131523; PubMed=8975721;
  RA Straub M.R., Natarajan D., Taylor L.B., Choi C.P., Martinie C.,
  RA Perbal B., Schofield P.N., Boulter C.A.;
  RT "Genomic structure and chromosomal mapping of the mouse nov gene.";
  RL Genomics 38:425-428(1996).
  [2]
  SEQUENCE FROM N.A.
  STRAIN=C57BL/6;
  RC MEDLINE=96204003; PubMed=8622864;
  RA Martinie C., Chevallier G., Rauscher F.J. III, Perbal B.;
  RT "Regulation of nov by WT1: a potential role for nov in
  RT nephrogenesis.";
  RL Oncogene 12:1479-1492(1996).
  CC -|- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
  CC GROWTH REGULATION (BY SIMILARITY).
  CC -|- SUBUNIT: Interacts with FBLN1 (By similarity).
  CC -|- SUBCELLULAR LOCATION: Secreted.
  CC -|- SIMILARITY: Contains 1 IGFBP domain.
  CC -|- SIMILARITY: Contains 1 WFCC domain.
  CC -|- SIMILARITY: Contains 1 TSP type-1 domain.
  CC -|- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
  CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
  CC the European Bioinformatics Institute. There are no restrictions on its
  CC use by non-profit institutions as long as its content is in no way
  CC modified and this statement is not removed. Usage by and for commercial
  CC entities requires a license agreement (See http://www.1st-sib.ch/announce/
  CC or send an email to license@1st-sib.ch).
  CC EMBL: X97863; CAA66457.1; -
  DR EMBL: Y09257; CAA70454.1; -
  DR EMBL: X96585; CAA65404.1; -
  DR MGD: MGI:109185; NOV.
  DR InterPro: IPR006208; Cys_knot.
  DR InterPro: IPR006207; Cys_knot_C.

```

```

DR InterPro: IPR000867; Ins1_gro_fac-pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; WFCC.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; TSP_1; 1.
DR Pfam: PF00093; WFCC; 1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; WFCC; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00223; IGF_BINDING; 1.
DR PROSITE: PS50092; TSP1; 1.
DR PROSITE: PS01208; WFCC_1; 1.
DR PROSITE: PS50184; WFCC_2; 1.
DR Growth factor; Signal.
KW SIGNAL
FT SIGNAL 1 21
FT CHAIN 22 354
FT DOMAIN 29 101
FT DOMAIN 102 168
FT DOMAIN 202 247
FT DOMAIN 261 335
FT DISULFID 261 298
FT DISULFID 278 312
FT DISULFID 289 328
FT DISULFID 292 330
FT DISULFID 297 334
FT CARBOHYD 91 91
FT CARBOHYD 277 277
SQ SEQUENCE 354 AA; 38928 MW; 08CE8C6C67829DE CRC64;

Query Match 39.38; Score 827; DB 1; Length 354;
Best Local Similarity 44.28; Pred. NO. 4.8e-54;
Matches 161; Conservative 51; Mismatches 114; Indels 38; Gaps 9;

QY 7 RTLAVALNTLH-LTRALS--TCPAC--HGPLAPKAPGAGVLRDGGCKYAKQNE 62
Db 8 RCLGLGLPHLLSIVASASLRCPSPKCPSPISPTCAPGRVSLDSCCPVCAKORGE 67
QY 63 DCSKTOPCDHTKGLIECNFGASSTALKGICRAOSEGRPEYNSRIYQNGESFOPNCKRQCT 122
Db 68 SCSEMRPPDSSGSLVCDRSADPNNGEIGICMV-PEGDNCVFDGVIYRNGEKIPEPCQYFCT 126
QY 123 CIDGAVGCTPLCPQELSLPMLGCPNRLKVSAGCCCEEWCE--DEDSIKSLDDQDLGL 180
Db 127 CRDQIGCLPRLCQLDVLPPDPCCPAPKKAAYVGECEKWTGSDQGTGTLG----LA 182
QY 181 LDASEVLETRNNELIAGKSSLRKLPVFGTEPRVLEFNPLAHGOKCIYVOTISQSCSKS 240
Db 183 LPVPRATVGEV----SDSI------NCTEQTTIMKASCSKSGCMG 216
QY 241 CGGISSTRVNDNDECRVLYETRICEVPRGQ-PVYSLKKKKCKSKTKSPFVRFTYAGCSS 299
Db 217 CGMGVSTRVTRNRLQCEWVQKTRICVPRCEQEPGEATDMKGGKCLRTKSKLSIHLQF 276
QY 300 GCSYVKKRYKCGSCVDGRCTPLQTRIVKMRRCEDGEMFSKNVMQISGICNYNCPHPNEA 359
Db 277 NCTSLYTKPRFCIGICSDGRCTPFNRTKIQVEFQCLPQGIKKRPVAVIGTCVCHSNCPQ 336
QY 360 PNEA 363
Db 337 NNEA 340

RESULT 15
NOV_XENLA STANDARD; PRT; 343 AA.
ID NOV_XENLA
AC P51609;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:02:31; Search time 32.913? Seconds

(without alignments)
2971.515 Million cell updates/sec

Title: US-09-495-448a-2

Perfect score: 2103

Sequence: 1 MSSSTFRTLAVALTLHLTR.....PNEASFRLYSLFNDIHKFRD 379

Scoring table: BIOSUM62

Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SPREMBL_23:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organelle:*
10: sp.phage:*
11: sp.plant:*
12: sp.podent:*
13: sp.virus:*
14: sp.vertibrate:*
15: sp.unclassified:*
16: sp.virus:*
17: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2076	98.7	379	11 Q9WTM9	Q5TM9 ratius norv
2	1624.5	77.2	375	13 Q98TX5	Q58TX xenopus lae
3	1455	69.2	334	4 Q9UID7	Q9UID7 homo sapien
4	938	44.6	344	13 Q98T08	Q58T8 gallus gall
5	937	44.6	343	13 Q42607	Q42607 xenopus lae
6	920.5	43.8	347	13 Q9PT80	Q5PT80 notophthalm
7	891	42.4	349	6 Q97765	Q97765 sus scrofa
8	827	39.3	354	11 Q8CA67	Q8CA67 mus musculu
9	773	36.8	367	4 Q95388	Q95388 homo sapien
10	758.5	36.1	367	11 Q54775	Q54775 mus musculu
11	757	36.0	367	11 Q99P00	Q99P00 ratius norv
12	635.5	30.2	331	4 Q95958	Q95958 homo sapien
13	622	29.6	334	4 Q95389	Q95389 homo sapien
14	518.5	24.7	251	11 Q8CIC8	Q8CIC8 mus musculu
15	499.5	23.8	280	4 Q9HCS3	Q9HCS3 homo sapien
16	495.5	23.6	176	13 Q9PSS6	Q9PSS6 gallus gall

17	374.5	17.8	128	11 Q9R2C0	Q9R2C0 ratius norv
18	345	16.4	113	11 Q92164	Q92164 ratius norv
19	343	16.3	119	11 Q91V29	Q91V29 mus musculu
20	343	16.3	119	11 Q920W6	Q920W6 mus spicile
21	341	16.2	100	4 Q9UDL6	Q9UDL6 homo sapien
22	298.5	14.2	374	5 Q9VVR0	Q9VVR0 drosophila
23	296	14.1	470	5 Q9VVK3	Q9VVK3 drosophila
24	293	13.9	230	4 Q8VYK7	Q8VYK7 homo sapien
25	195	9.3	77	4 Q9UDE4	Q9UDE4 homo sapien
26	184	8.7	70	13 Q9DF21	Q9DF21 scyllorhinu
27	179	8.5	2327	13 Q9IBG7	Q9IBG7 xenopus lae
28	171	8.1	1045	5 Q8T3A6	Q8T3A6 caenorhabdi
29	171	8.1	1070	5 Q8T3A7	Q8T3A7 caenorhabdi
30	171	8.1	1111	5 Q9XMD6	Q9XMD6 caenorhabdi
31	167.5	8.0	792	13 Q90Z43	Q90Z43 gallus gall
32	167.5	8.0	1095	13 Q90XG4	Q90XG4 gallus gall
33	167	7.9	58	6 Q97574	Q97574 bos taurus
34	166	7.9	1034	11 Q35888	Q35888 ratius norv
35	166	7.9	1521	4 Q95710	Q95710 homo sapien
36	166	7.9	1525	4 Q9Y5Q7	Q9Y5Q7 homo sapien
37	166	7.9	1529	4 Q94813	Q94813 homo sapien
38	165.5	7.9	950	13 Q90Z44	Q90Z44 gallus gall
39	163	7.8	1036	4 Q9NZV1	Q9NZV1 homo sapien
40	162	7.7	360	5 Q8MN64	Q8MN64 dictyosteli
41	161.5	7.7	406	11 Q92513	Q92513 mus musculu
42	161.5	7.7	426	11 Q8VBA6	Q8VBA6 mus musculu
43	160.5	7.6	2104	5 Q21281	Q21281 caenorhabdi
44	160.5	7.6	2104	5 Q964N4	Q964N4 caenorhabdi
45	158.5	7.5	1028	11 Q9JLL0	Q9JLL0 mus musculu

ALIGNMENTS

RESULT 1

ID	Q9WTM9	PRELIMINARY:	PRT:	379 AA.
AC	Q9WTM9			
DT	01-NOV-1999 (TREMBLrel. 12, Created)			
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	CYR61 precursor.			
GN	CYR61.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
KC	STRAIN-12m; TISSUE-Aorta;			
RA	Unoki H., Yonekura H., Furukawa K., Yamamoto H.;			
RT	"Rat Cyr61 mRNA."			
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AB015877; BAA78339.1; -			
DR	InterPro: IPR006208; Cys_knot.			
DR	InterPro: IPR006207; Cys_knot_C.			
DR	InterPro: IPR000867; Insl_gro_fac.pr.			
DR	InterPro: IPR000884; TSPL.			
DR	InterPro: IPR001007; VWF_C.			
DR	Pfam: PF00007; Cys_knot; 1.			
DR	Pfam: PF00219; IGFBP; 1.			
DR	Pfam: PF00090; tsp_1; 1.			
DR	Pfam: PF00093; vwc; 1.			
DR	SMART: SM00041; CT; 1.			
DR	SMART: SM00121; IB; 1.			
DR	SMART: SM00209; TSPL; 1.			
DR	SMART: SM00214; VWC; 1.			
DR	PROSITE: PS01185; CTCK_1; 1.			
DR	PROSITE: PS01225; CTCK_2; 1.			
DR	PROSITE: PS00222; IGF_BINDING; 1.			
DR	PROSITE: PS00093; TSPL; 1.			
DR	PROSITE: PS01208; VWF; 1.			
KW	Signal.			

FT SIGNAL 1 24 POTENTIAL.
CYR61.
SQ SEQUENCE 379 AA: 41728 MW: D2ABAF077B84762B CRC64:

Query Match 98.7%; Score 2076; DB 11; Length 379;
Best Local Similarity 98.2%; Pred. No. 1.8e-189;
Matches 372; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSSSTRTAVANTLTHTLRLALSTCPACCHCPLEAPKCAPGVGLVBDGCGCCVCAKOL 60
DB 1 MSSSTRTAVANTLTHTLRLALSTCPACCHCPLEAPKCAPGVGLVBDGCGCCVCAKOL 60
QY 61 NEDCKSTQPCDHRTKGLCECNFASSTALKGICRAQSEGRPEYNSRIYQNGSEFQPNCKHQ 120
DB 61 NEDCKSTQPCDHRTKGLCECNFASSTALKGICRAQSEGRPEYNSRIYQNGSEFQPNCKHQ 120
QY 121 CTCIDGAVGICPLCPPELSPNLGCPNPLVYKSGCCCEWVCEDESIKSLDDDDLLG 180
DB 121 CTCIDGAVGICPLCPPELSPNLGCPNPLVYKSGCCCEWVCEDESIKSLDDDDLLG 180
QY 181 LDASEVELTRNNELIAGKSSLRPLVEGTEPRVLNPLHAHQKCIYQTTSMSCSKS 240
DB 181 LDASEVELTRNNELIAGKSSLRPLVEGTEPRVLNPLHAHQKCIYQTTSMSCSKS 240
QY 241 CGTGISTRTYNDNPECHLVETRICVRRPGQPYVSSLRKGGKSKTKSPPEVRYTAG 300
DB 241 CGTGISTRTYNDNPECHLVETRICVRRPGQPYVSSLRKGGKSKTKSPPEVRYTAG 300
QY 301 GSSVKKYRPKYCGSCVDGRCCTPLQRTYKMRRCEDGEMFSKNVMIIOSCKNVCNCPHP 360
DB 301 GSSVKKYRPKYCGSCVDGRCCTPLQRTYKMRRCEDGEMFSKNVMIIOSCKNVCNCPHP 360
QY 361 NEASFRLYSLFNDIHKFRD 379
DB 361 NEASFRLYSLFNDIHKFRD 379

RESULT 2

Q98TX5 PRELIMINARY; PRT; 375 AA.
AC 098TX5;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Secreted cysteine-rich protein cyr61.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodidae; Xenopus.
NCBI_TaxID=6353;
RN NCBI_TaxID=6353;
RP SEQUENCE FROM N.A.
RA Latinkic B.V., Bennett B., Smith J.C.;
RT "Characterization of Xenopus cyr61."
RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
EMBL: AF20592; AAK00947.1;
DR EMBL: AF20592; AAK00947.1;
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot.C.
DR InterPro: IPR000867; Insl_gro_fac-pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; VWF.C.
DR Pfam: PF00007; Cys_knot.1.
DR Pfam: PF00090; IGFBP.1.
DR Pfam: PF00093; vwc.1.
DR SMART: SM00041; CT.1.
DR SMART: SM00121; IB.1.
DR SMART: SM00209; TSP1.1.
DR SMART: SM00214; VWC.1.
DR PROSITE: PS01185; CTCK.1; 1.
DR PROSITE: PS01225; CTCK.2; 1.
DR PROSITE: PS00222; IGF_BINDING.1.
DR PROSITE: PS0092; TSP1.1.

DR PROSITE: PS01208; VWF.C.1.
SQ SEQUENCE 375 AA: 41460 MW: 78075CA7B380304E CRC64:

Query Match 77.2%; Score 1624.5; DB 13; Length 375;
Best Local Similarity 77.5%; Pred. No. 1.9e-146;
Matches 293; Conservative 29; Mismatches 37; Indels 19; Gaps 7;

QY 9 LAVALTLHLTRTALSTCPACCHCPLEAPKCAPGVGLVBDGCGCCVCAKOLNEDCKSTQ 68
DB 10 LAVAL-TLGFIDLAVSSCPAVGOCPEYKCAPGVGLVBDGCGCCVCAKOLNEDCKSTQ 68
QY 69 PCDHRTKGLCECNFASSTALKGICRAQSEGRPEYNSRIYQNGSEFQPNCKHCTCIDGAV 128
DB 69 PCDHRTKGLCECNFASSTALKGICRAQSEGRPEYNSRIYQNGSEFQPNCKHCTCIDGAV 128
QY 129 GCPLCPPELSPNLGCPNPLVYKSGCCCEWVCEDESIKSLDDDDLLG 184
DB 129 GCPLCPPELSPNLGCPNPLVYKSGCCCEWVCEDESIKSLDDDDLLG 184
QY 185 EVELTRNNELIAGKSSLRPLVEGTEPRVLNPLHAHQKCIYQTTSMSCSKS 241
DB 187 EGELTRKNEFAVING-GIKMLPVGSDPQ-----SHVENSKCIYQTTSMSCSKS 238
QY 242 GTGISTRTYNDNPECHLVETRICVRRPGQPYVSSLRKGGKSKTKSPPEVRYTAG 301
DB 239 GTGISTRTYNDNPECHLVETRICVRRPGQPYVSSLRKGGKSKTKSPPEVRYTAG 298
QY 302 GSSVKKYRPKYCGSCVDGRCCTPLQRTYKMRRCEDGEMFSKNVMIIOSCKNVCNCPHP 361
DB 299 GSSVKKYRPKYCGSCVDGRCCTPLQRTYKMRRCEDGEMFSKNVMIIOSCKNVCNCPHP 358
QY 362 EASFRLYSLFNDIHKFRD 379
DB 359 EA-YPYRRLFNDIHKFRD 375

RESULT 3

Q9UID7 PRELIMINARY; PRT; 334 AA.
AC 09UID7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Cyr61 protein.
GN CYR61.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN NCBI_TaxID=9606;
RP SEQUENCE OF 1-107 FROM N.A.
RA Anding B., Long Y.;
RT TISSUE-Kidney;
RT "Cloning of a new gene down-regulated in the small-cell tumor embryonal-rhabdomyosarcoma (RMS)."
RT Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
EMBL: AF003114; AAF21597.1;
DR EMBL: AF003114; AAF21597.1;
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot.C.
DR InterPro: IPR000867; Insl_gro_fac-pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; VWF.C.
DR Pfam: PF00007; Cys_knot.1.
DR Pfam: PF00090; IGFBP.1.
DR Pfam: PF00093; vwc.1.
DR SMART: SM00041; CT.1.
DR SMART: SM00209; TSP1.1.
DR SMART: SM00214; VWC.1.
DR PROSITE: PS01225; CTCK.2; 1.
DR PROSITE: PS01208; VWF.C.1.
DR PROSITE: PS01208; VWF.C.1.
SQ SEQUENCE 334 AA: 37246 MW: 918987A7A352E948 CRC64:

	Query Match	69.2%	Score 1455	DB 4	Length 314
	Similarity	78.9%	Pred. No. 2.3e-130		
	Match	Local	Conservative	13	Mismatches 32
					Indels 28
					Gaps 6
QY	49	GCGCCVCAKOLNEDCSKTOPCDHFKGLGECNFGASSTLAKIGICRAOSEIRPEYNSRIQO	108		
Db	2	GCGRHNPNCILHGHTASPTSYKHHKTKGLGECNFGASSTLAKIGICRAOSEIRPEYNSRIQO	61		
QY	109	NGESFPQNCKHQCTCIDGAVGCIPLCPQELSLPNIGCPNRLVYKSGQICEEWDCEDSI	168		
Db	62	NGESFQPNCKHQCTCIDGAVGCIPLCPQELSLPNIGCPNRLVYKSGQICEEWDCEDSI	121		
QY	169	KDSLDODD-----IGLDAFVELTRNNELLAIGKSSSLKRLPVGTEIRVLFPNLHANG	224		
		:			
Db	122	KDPMDDOGLGKELGKDFPDAFSEVLRNNELLAVGSSSLKRLPVGMEIRIYNPL--OG	179		
QY	225	QKCIYQTTWSQCSKSCGTGISTRYTNNPPECRLKETRIFICVPRCQGVYSSLKKGKRC	284		
Db	180	QKCIYQTTWSQCSKSCGTGISTRYTNNPPECRLKETRIFICVPRCQGVYSSLKKGKRC	239		
QY	285	SKTKKSPVPVFTYTAGSSVYKRYPKYCGSCYDVGKCCPLQTRVYKMRICED-----	337		
Db	240	SKTKKSPVPVFTYTAGCLSYKRYPKYCGSCYDVGKCCPLQTRT-----CEDAVPLRRM	293		
QY	338	GEMSKN-----YMMIQSCKMYNCPHPPEASRLATSLNDNHKFD	379		
Db	294	GDIQERHDEPL-----KCNINCNCHAAEAFPEYRLNDNHKFD	334		

AC	ID	098T08	PRELIMINARY;	PRT;	344 AA.
DT	01-JUN-2001	(TREMBLrel. 17, Created)			
DT	01-JUN-2001	(TREMBLrel. 17, Last sequence update)			
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)			
DE	Connective tissue growth factor precursor (Connective tissue growth factor/hypertrophic chondrocyte-specific protein 24).				
OS	Gallus gallus (Chicken).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;				
OC	Gallus.				
XX	NCBI_TaxID=9031;				
RP	[1]				
RA	SEQUENCE FROM N.A.				
RL	Gygi D.;				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AJ298335; CAC33458.1; -				
DR	EMBL; AF463517; AAL68834.1; -				
DR	InterPro: IPR006208; Cys_Knot.				
DR	InterPro: IPR006207; Cys_Knot_C.				
DR	InterPro: IPR000867; Insl_gro_fac_pr.				
DR	InterPro: IPR001007; TSP1.				
DR	InterPro: IPR001007; TSP1.				
DR	Pfam; PF00207; Cys_knot; 1.				
DR	Pfam; PF00219; IGFBP; 1.				
DR	Pfam; PF00090; tsp_1; 1.				
DR	Pfam; PF00093; vwc; 1.				
DR	SMART; SM00041; CT; 1.				
DR	SMART; SM00121; IB; 1.				
DR	SMART; SM00209; TSP1; 1.				
DR	SMART; SM00214; vwc; 1.				
DR	PROSITE: PS01185; CTCK_1; 1.				
DR	PROSITE: PS01225; CTCK_2; 1.				
DR	PROSITE: PS00222; TGF_BINDING; 1.				
DR	PROSITE: PS00092; TSP1; 1.				
DR	PROSITE: PS01208; vwc; 1.				
DR	Signal.				

FT	SIGNAL	1	21	POTENTIAL.
CH	CHAIN	22	344	CONNECTIVE TISSUE GROWTH FACTOR.
SEQ	SEQUENCE	344 AA:	37499 MW:	69E639AF6B9BFD00 CR664;
	Query Match	44.6%;	Score 938;	DB 13;
	Best Local Similarity	45.7%;	Pred. No. 4,4e-81;	
	Matches 170:	Conservative	61;	Mismatches 103; Indels 38; Gaps 7
QY		8 TLAAVAVTLHLH-TLALSTCPAACGCP-L-EAPKCAPGVGLVADGGCGCCVYCAKQINDECS	65	
Db		5 SLAAVALLALGLPEVRGEGESGQCCGCGSGPGSGPAGVSLYLDGGCCRCVCAKQIGELCT	64	
QY		66 KTOPEDHTKGLTENCNGASTALGKICIRAOSESPREYMSIRYQNGESEFOPNKHQCTCID	125	
Db		65 ERDPDHHKGLFCDGSGPANRRIRIGCTAR-DSAPCVFSGMYRSESPQSSCKYQCTCID	123	
QY		126 GAVGCIPLCPQBELSLPNIGCPNPLRYKVSQGCCBEMWVCDENSIKSLDQDDDLGLDASE	185	
Db		124 GAVGCVPLCSMDVIRLPSDPCPYPRRYVKLPKGCBEWVCD-----	163	
QY		186 VELTRNNNELIAGKSSSLKL-RPVGTGEPRVLFNPLHAHGOKCIQVOTSMSCSKSCGTG	244	
Db		164 -----AKQGTAVGPALAAVRLIEDTYGPRPTM-----RANCLVQTTTMSKSKTCCMG	211	
QY		245 ISTRVTDNPECRILVETRIICEVRDGGQPVYSSLLAKGKKSKTKSPPEVRFVTIAGCQSV	304	
Db		212 ISTRVTDNNAFCRLTEKOSRCLMWRCDEADLENIKKGKKCIPTPKISIKFELSGCTSY	271	
QY		305 KKRYRKVYGSQVDCGCTPLQTRVYKMRRCGDEGEMFNSNMIMOSCRKNYNCPPHNEMAS	364	
Db		272 KTYRKAFGCVGCTDGGCTPHRTALRVEKRCADGSELTKRMKMFITTCACHYNCQGNDTF	331	

```

RESULT 5
042607
ID 042607 PRELIMINARY; PRT; 343 AA:
AC 042607;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, last annotation update)
DE Connective tissue growth factor XCTGF.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Ying Z., King M.L.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U43524; AAB67638.1; +
DR EMBL; U43523; AAB67638.1; +
DR InterPro: IPR0066208; Cys_knot.
DR InterPro: IPR0066207; Cys_knot_C.
DR InterPro: IPR000867; Ins1_gro_fac.pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; WFC.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00219; IGFBP; 1.
DR Pfam; PF00090; tsp_1; 1.
DR Pfam; PF00093; WFC; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; WFC; 1.
DR PROSITE; PS01185; CMCK_1; 1.
DR PROSITE; PS01225; CMCK_2; 1.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS01208; WFC; 1.

```

SO	SEQUENCE	343 AA;	37966 MW;	9392221C5DB565A81	CRC64;
Query Match		44.6%;	Score 937;	DB 13;	Length 343;
Best Local Similarity		46.7%;	Pred. No. 5,5e-81;		
Matches	165;	Conservative	60;	Mismatches	90;
				Indels	38;
				Gaps	6
OY	26	CPAACHCHLEAPKAPAGVGLVRDGGCCCKVKVAKOLNEQCSKTPQDPDHHKGLGECNFGASST	85		
Db	24	CNGBEQCCNNKVPVCDPGBVMVDGGCCCKVKCSKQJGELCTENDVCDPHGGLGCDGSGRYN	83		
OY	86	ALKGICRAOSEGRPEYNSRIYONSESOPNNHOCCTIDGAVGICPLCPQELSPNLGIC	145		
Db	84	RKIGVCTAR--BGAPCVFGGTYVRSESPQSSCKYCTCIDGGVGCVPILCSMDIRLPSPEC	142		
OY	146	PNPRLVYKVSQCCCEWVCDDEDSIKDSLDDQDLG--LDASVELTRNNELIATIGKSSL	203		
Db	143	PPPRVVKLPGGCCCEWVCDQ--DQERTLVPAIPAFMEET-----	181		
OY	204	KRLPFGEGPRLVFLPRLAHAGQKCIYQFTSMQCSKSGTGTSTRYTNDNPECRLVKTR	263		
Db	182	-----YGDPSLI-----RANCLVQTTTMSACSKTGMMGISTRYTNDNECRLEKQSR	229		
OY	264	ICEVAPCGQPYVSLITKGGKCSKTKKSPBPVATFYAGSSVYKRYPRKYCGSCVDRGCTP	323		
Db	230	LCMVRPCPADLEENIKKIKKCCIRPKIKSPVAFEESSGCTSVXYTAKFCGVCYTDGRCTP	289		
OY	324	LQRTVYKRRFCEDEGEMFSKRVNMAQSCKANTNCHPNVDSRLR--SLFNDI	374		
Db	290	HRITATLPEYFECPCDGEVYKKNMMFLKTCACHENCGADDIFBAMVYRKMYGDM	342		

RESULT 6			
ID	NAME	PRELIMINARY	PRT; 347 AA.
Q9P780	Q9P780		
AC	Q9P780;		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, last sequence update)		
DT	01-MAR-2003 (TREMBLrel. 23, last annotation update)		
DE	Connective tissue growth factor.		
GN	CTGF.		
OS	Notophthalmus viridescens (Eastern newt) (Triturus viridescens).		
OC	Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Euphylaria; Batrachia; Caudata; Salamandridae; Salamandridae;		
OC	Notophthalmus.		
OX	NCBI_TaxID=8316;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Forelimb blastema;		
RA	Gates P.B.;		
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Forelimb blastema;		
RX	MEDLINE=99033008; PubMed=9813273;		
RX	Cash D.E., Gates P.B., Imokawa Y., Brookes J.P.;		
RA	retinoid regulation in limb blastemal cells.";		
RT	"Identification of new connective tissue growth factor as a target of		
RL	Gene 222:119-124(1998).		
RL	EMBL; AJ271167; CAB65965.1; -		
DR	InterPro: IPR006208; Cys_knot.		
DR	InterPro: IPR006207; Cys_knot.C.		
DR	InterPro: IPR000867; Insl_gro_fac.pr.		
DR	InterPro: IPR000884; TSP1.		
DR	InterPro: IPR001007; WVF.C.		
DR	Pfam: PF00007; Cys_knot; 1.		
DR	Pfam: PF00219; IGFBP; 1.		
DR	Pfam: PF00090; tsp_1; 1.		
DR	Pfam: PF00093; WVC; 1.		
DR	SMART: SM00041; CT; 1.		
DR	SMART: SM00121; IB; 1.		
DR	SMART: SM00209; TSP1; 1.		
DR	SMART: SM00214; WVC; 1.		
DR	PROSITE: PS01185; CTCK_1; 1.		

QY	DR	PROSITE: PS01225; CTRK 2; 1.
DR	PROSITE: PS00222; TGF_BINDING; 1.	
DR	PROSITE: PS50092; TSP1; 1.	
DR	PROSITE: PS01208; VWFC; 1.	
SQ	SEQUENCE 347 AA; 38098 MM; 387E2399F27672C1 CRC64;	
Query Match	43.8%; Score 920.5; DB 13; Length 347;	
Best Local Similarity	45.1%; Pred. No. 2.1e-79;	
Matches 169; Conservative	54; Mismatches 103; Indels 49; Gaps 6	
QY	9 LAVAVTTLHLTRLALSTCPACHPCHLEAPKCAPGVGLVRDGGCGCKVCAKQNLNEDCSKTQ 68	
DB	12 LLIAVALLSNWSCA-QDCSGECRCNKRPECPAGSLVMWDGCGCKVCAKQGLGELCTEED 70	
QY	69 PCDHTKGLGECNFGASSTALKGICRAQSGRPCEYSRIYQNGESFQPNCKHQCCTIDAV 128	
DB	71 VCDPHRGLFCDFGSRVNNKIGVCTAK-DGAPCVFGMAYRSGSEFSQCKYQCTCIDGCV 129	
QY	129 GCIPLCPELSPNLGCPNPRLYKSGGCGCEMWYCDDEDSIKSLDDQDLDLGLDASEVFL 188	
DB	130 GCVPLCGADVRLPSPDCFPFRVKQGLGKCEEWACDQPK-----EQ 170	
QY	189 TRNNELIY-----GKSKSLRLRPVGTBPRVLFNPLAHGQKCIYVOTTSMSQSKSC 241	
DB	171 TRVGFALVYRQEEFYGPDSILMR-----ANCLVOTTSMSQSKSC 211	
QY	242 GTGISTRTVTDNPECRILVETRIKICEVRPGQGVVSYSLKKGKCSKTKSPPEVREFTYAC 301	
DB	212 GMGISTRTVTDNENMOKRLKQSLCMLNVRCEADLEENIKKGCICIRPKISRVKFLSLSC 271	
QY	302 SSVKKYRPRKYCGSGCVDRGQCPPLQTRIVYKMRFCEDGEMFSKNVMIOQSCKNVNCAPHN 361	
DB	272 TSVKIRAFKFGCVCTDRCGCPRHPRATPLRVEKFCEDGEVKKMKMFIRKTCACHYNCQGN 331	
QY	362 EASFRLY--SLFNDI 374	
DB	332 DIFESMYRKMYGDM 346	

RESULT	7			
ID	097765	PRELIMINARY:	PRT:	349 AA.
AC	097765;			
DT	01-MAY-1999 (TrEMBLrel. 10, Created)			
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)			
DE	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)			
GN	Connective tissue growth factor.			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RA	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Uterus;			
RA	Harding P.A., Brigsstock D.R.;			
RT	"Cloning and sequencing of a porcine connective tissue growth factor			
RL	(cDNA)." ;			
RL	Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; U700060; AAD00174.1; -.			
DR	InterPro: IPR006208; Cys_knot.			
DR	InterPro: IPR006207; Cys_knot.C.			
DR	InterPro: IPR000867; InsL_gro_fac.pr.			
DR	InterPro: IPR000884; TSP1.			
DR	InterPro: IPR001007; WF_C.			
DR	Pfam; PF00007; Cys_knot.1.			
DR	Pfam; PF00219; IGFBP.1.			
DR	Pfam; PF00090; tsp.1; 1.			
DR	Pfam; PF00093; wvc.1.			
DR	SMART; SM00041; CT; 1.			
DR	SMART; SM00121; IB; 1.			
DR	SMART; SM00209; TSP1.1.			
DR	SMART; SM00214; WVC; 1.			

DR PROSITE; PS0092; TSP1; 1.
DR PROSITE; PS01208; VWF; 1.
SO SEQUENCE 367 AA; 40331 MW; 9F29CA94D69C0502 CRC64;
Query Match 36.8%; Score 773; DB 4; Length 367;
Best Local Similarity 38.4%; Pred. No. 2,5e-65;
Matches 146; Conservative 58; Mismatches 124; Indels 52; Gaps 7;

9 LAVAVTLHLTRALST-----CPAACHCPLEAPKCAPGVGLVRDGGCCGCAKOLN 61
25 LSPAPTTMDTPAPLEDTSSRFQCKWPCPCPPSPRCPLGVSLLTDGCCCCMACQOLG 84
62 EDCSKTPCDHTKGLCNF--GASSTALKGICRAOSEGRPCVNSRIYONGESFQPNCKH 120
85 DNGTEALCDPHRGVLCYDYGSDRPRAIGVC--AQVGVGCVLDGVRRNNGSQSPQNKYN 143
121 CTCIDGAVGCIPLCPQELSLPNLGICRPRPVKYSGGCCCEWVCDSDSKDSDDDLLG 180
144 CTCIDGAVGCTPLC--LRVRPRLMCPHPRVSIIPGCCCEWVEDAKPRKTAAPRDGTA 202
181 LDA-SEVELTRNNELIAGKSSLKRLPVGTEPRVLFNPLAHGOKCIQVOTSMQSCSK 239
203 FDAVGEVEMHNR-----CIATYSPMSPCST 228
240 SCGTGISTRYTNDNPECRLVETRICVRCGQPVYSSLKGRKCKTKKSPBPVRYTA 299
229 SCGLGVSTRISNNACWPEQESRLCNLRPCVDVIDLITKAGKCKLAVYQPEASMFLLA 288
300 GCSVVKYRKYKCGSVDCRCCTPLQTRVYKMFREDEGEMFSKNMIMOSCKNCPH 359
289 GCVSTSTYRKYKCGVDCNRCCIPYKSKTIVDFQCPDGLGFSROVLWLNACGNCNCRN 348
360 PNEASFRLYSLFNDIHKFRD 379
349 PND-----IFADLESYPD 361

RESULT 10
054775 PRELIMINARY; PRT; 367 AA.
AC 054775;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE ELM1.
GN ELM1 OR WISPI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HEN;
RX MEDLINE=98119879; PubMed=9449709;
RA Hashimoto Y., Shindo-Okada N., Tani M., Nagamachi Y., Takeuchi K.,
RA Shiroishi T., Toma H., Yokota J.;
RT "Expression of the ELM1 gene, a novel gene of the CCN (connective
RT tissue growth factor, Cyr61/Ce10, and neuroblastoma overexpressed
RT gene) family, suppresses in vivo tumor growth and metastasis of K-1735
RT murine melanoma cells.";
RL J. Exp. Med. 187:289-296(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=99061933; PubMed=9843955;
RA Penica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,
RA Brush J., Taneyhill L.A., Denel B., Lew M., Matanabe C., Cohen R.L.,
RA Melham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.D.,
RA Gurney A.L., Botstein D., Levine A.J.;
RT "Wisp genes are members of the connective tissue growth factor family
RT that are up-regulated in wnt-1-transformed cells and aberrantly
RT expressed in human colon tumors.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).

DR EMBL; AB004873; BAA24949.1; -
DR EMBL; AF100777; AAC69319.1; -
DR MGD; MGI:1197008; Wisp1.
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR000867; Insl_gro_fac_pr.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; VWF_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00219; IGFBR; 1.
DR Pfam; PF00090; LSP_1; 1.
DR Pfam; PF00093; VWF; 1.
DR SMART; SM00041; CT; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VWF; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS0092; TSP1; 1.
DR PROSITE; PS01208; VWF; 1.
SO SEQUENCE 367 AA; 40702 MW; 3B7C0569EFAB5E96 CRC64;

Query Match 36.1%; Score 758.5; DB 11; Length 367;
Best Local Similarity 37.6%; Pred. No. 6e-64;
Matches 143; Conservative 55; Mismatches 135; Indels 47; Gaps 6;

3 SSTRFLAVAVTLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCGCAKOLN 62
26 SPTPTMTTPAPLETTTRPECKKPCPCSPRCPLGVSLLTDGCCCKCAQOLG 85
63 DCSKTPCDHTKGLCNF--GASSTALKGICRAOSEGRPCVNSRIYONGESFQPNCKH 121
86 NCTEALCDPHRGVLCYDYGSDRPRAIGVC--AQVGVGCVLDGVRRNNGSQSPQNKYN 144
122 TCIDGAVGCIPLC--PQELSLPNLGICRPRPVKYSGGCCCEWVCDSDSKDSDDDLL 179
145 TCIDGAVGCTPLDLSRP---PLMCRQPRHVRVPGCCCEWVEDAKPRKTAAPRDGTA 192
180 GLDASEVELTRNNELIAGKSSLKRLPVGTEPRVLFNPLAHGOKCIQVOTSMQSCSK 239
193 -----RQTALDLDRFAASGAVGQRENCIATYSPMSPCST 228
240 SCGTGISTRYTNDNPECRLVETRICVRCGQPVYSSLKGRKCKTKKSPBPVRYTA 299
229 TCGIGISTRIISNNACWPEQESRLCNLRPCVDVIDLITKAGKCKLAVYQPEATNFTLA 288
300 GCSVVKYRKYKCGSVDCRCCTPLQTRVYKMFREDEGEMFSKNMIMOSCKNCPH 359
289 GCVSTSTYRKYKCGVDCNRCCIPYKSKTIVDFQCPDGLGFSROVLWLNACGNCNCRN 348
360 PNEASFRLYSLFNDIHKFRD 379
349 PND-----IFADLESYPD 361

RESULT 11
099P0 PRELIMINARY; PRT; 367 AA.
AC 099P0;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE ELM1.
GN ELM1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20487548; PubMed=11031104;
RA Sleeman M.A., Murison J.G., Strachan L., Kumble K.D., Glenn M.P.,
RA McGath A., Bickerstaff P., Grierson A., Havukkala I., Tan P.,

RX MEDLINE=99061933; PubMed=9843955;
 RA Penica D., Swenson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,
 RA Brush J., Taneyhill L.A., Deuel B., Lew M., Matnabe C., Cohen R.L.,
 RA Welham M.F., Finley G.G., Quirke P., Goddard A.D., Hillan K.J.,
 RA Gurney A.L., Botstein D., Levine A.J.;
 RT "WIS genes are members of the connective tissue growth factor family
 RT that are up-regulated in wt-1-transformed cells and aberrantly
 RT expressed in human colon tumors."
 RL Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).
 DR EMBL: AF100781; AAC96323.1;
 DR Genbank: HGNC12771; WISP3.
 DR Interpro: IPR006208; Cys_knot.
 DR Interpro: IPR006207; Cys_knot_C.
 DR Interpro: IPR000867; Ins1_gro_fac_pr.
 DR Interpro: IPR000884; TSP1.
 DR Pfam: PF00007; Cys_knot.1.
 DR Pfam: PF00029; IGFBP.1.
 DR Pfam: PF00090; TSP.1; 1.
 DR SMART: SM00041; CT.1.
 DR SMART: SM00121; IB.1.
 DR SMART: SM00209; TSP1.1.
 DR PROSITE: PS01225; CTCK.2.1.
 DR PROSITE: PS00222; IGF-BINDING.1.
 DR PROSITE: PS50092; TSP1.1.
 DR PROSITE: PS50092; TSP1.1.
 SQ SEQUENCE 354 AA; 39292 MW; 67F48D0D5C2F5EE3 CRC64;

Query Match 29.6%; Score 622; DB 4; Length 354;

Best Local Similarity 34.4%; Pred. No. 5.9e-51;
Matches 118; Conservative 56; Mismatches 119; Indels 50; Gaps 8;

QY 26 CPACHCPLPAPKAPGVLVNDGCGCCVCAKOLNEDSKTOPCDHTGLEGCFGASST 85
 DB 48 CHMPCKPQKPCPGVSLVRDGGCCCKKCAKOPKICNEADLCDFHGLGYCDYGVDP 107
 QY 86 ALK-GICRAQSEKREKSRITONGESFOPNCKHCTCIDAVGICPLCPOLSLPNEG 144
 DB 108 RIETGVC-AYLVAVGSEFQVHNMQVQPNPLFSCILVSGAIGCTP-----LFTPKLA 161
 QY 145 ---CPNRLVYKVGCCCEWVDEDSIKSLDDODDLGLDASEVELTRNNELIAIGKS 201
 DB 162 GSHCSGAK---GGRKSDQSNCSLEPDLQO-----ST 190
 QY 202 SLKRLPVEGTPEPRVLEPNLHAHQKCIYVTTSSQCSKSCGICSTRVNDNPECKLYE 261
 DB 191 SKYTMPAYRNPLI-----WKKKCLVQATKWTPCSRICGMGINSRVNENSCMKRE 243
 QY 262 TRICEVRPGGOPYSLK--KGKCKSKTKKSEPRVETFAAGSSVKKYPRKCGGVDR 319
 DB 244 KRLCTIOPCDNSLTKIKIPKGTCTOPTQLSKAEKFEVSGCSSTOSYKPTFCGICLDR 303
 QY 320 CCTPLQTRVKKRFCEDEGEMFSKNYMTIQSCCKANYNCPHNE 362
 DB 304 CCIPNKSMTITQFDCPNESGFKMKMLWITSCVCQNCNREPQD 346

RESULT 14

ID 08CIC8 PRELIMINARY; PRT; 251 AA.
 AC 08CIC8:
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Similar to WNT1 Inducible signaling pathway protein 2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10900;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC032877; AAH32877.1; -

SQ SEQUENCE 251 AA; 27106 MW; 57520309A9069524 CRC64;

Query Match 24.7%; Score 518.5; DB 11; Length 251;
 Best Local Similarity 37.5%; Pred. No. 2.9e-41;
 Matches 100; Conservative 41; Mismatches 81; Indels 45; Gaps 5;

QY 15 LHLTLRLALS-----TCPACHCPLPAPKAPGVLVNDGCGCCVCAKOLNED 63
 DB 6 LHLTLRLALS-----TCPACHCPLPAPKAPGVLVNDGCGCCVCAKOLNED 63
 QY 64 CSKTOPCDHTKLEGCFGASSTALKICCAQSGRCCEVNSRTYONGESFOPNCKHCTC 123
 DB 64 CHLHVNCNPSQGVYCPGAGPSGRGVYCLFEEDGSEVNGRATYLDGETFKPCRYLRC 123
 QY 124 INGAVGICPLCPOLSLPNEGCPNRLVYKVGCCCEWVDEDSIKSLDDODDLGLDA 183
 DB 124 DGGFTCPDPLCSDDVLPMDCDPRRIQVPGRCCEWVCDQAVMPAL----- 172
 QY 184 SEVELTRNNELIAIGKSLKRLPVGTPEPRVLEPNLHAHQKCIYVTTSSQCSKSCGT 243
 DB 173 -----QPSAQ-----GHQLSLVTPASADG-PCPMWSTAMGPCSTTGL 211
 QY 244 GISTRVNDNPECKLYEETRICVPRC 270
 DB 212 GIATRVSNQNRFCQLEIQRRLCLSRPC 238

RESULT 15

ID 09HCS3 PRELIMINARY; PRT; 280 AA.
 AC 09HCS3:
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE WISP-1 variant.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tanaka S., Sugimachi K.;
 RT "Human WISP-1 variant."
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB034725; BAB17849.1; -
 DR Interpro: IPR006208; Cys_knot.
 DR Interpro: IPR006207; Cys_knot_C.
 DR Interpro: IPR000867; Ins1_gro_fac_pr.
 DR Interpro: IPR000884; TSP1.
 DR Pfam: PF00007; Cys_knot.1.
 DR Pfam: PF00029; IGFBP.1.
 DR Pfam: PF00090; TSP.1; 1.
 DR SMART: SM00041; CT.1.
 DR SMART: SM00121; IB.1.
 DR SMART: SM00209; TSP1.1.
 DR PROSITE: PS01185; CTCK.1.1.
 DR PROSITE: PS01225; CTCK.2.1.
 DR PROSITE: PS50092; TSP1.1.
 SQ SEQUENCE 280 AA; 30697 MW; 26B254D4A060738E CRC64;

Query Match 23.8%; Score 499.5; DB 4; Length 280;
 Best Local Similarity 28.6%; Pred. No. 2.1e-39;
 Matches 108; Conservative 46; Mismatches 89; Indels 135; Gaps 8;

QY 9 LAVAVTILHLTLRLALST-----CPACHCPLPAPKAPGVLVNDGCGCCVCAKOLN 61
 DB 25 LSPAPTTMDFTPAPELDTSSRPQFKWPCPCPPSPROPGLVSLITDGECCCKMAQOIG 84
 QY 62 EDCSKTOPCDHTKLEGCFGASSTALKICCAQSGRCCEVNSRTYONGESFOPNCKHCTC 121
 DB 85 DICTEALICDPHRLGLYCDY-----SGDRP-----RYAIGV-----CAH-- 117
 QY 122 TCIDGAVGICPLCPOLSLPNEGCPNRLVYKVGCCCEWVDEDSIKSLDDODDLGL 181


```
Db      118  ----AVG----- 120
QY      182  DASFEVLTRNNELAIIGKSSLKRLPVFGTEPRVLEPNLHAHGKCIIVOTISMSQCSKC 241
Db      121  ---EVEAMHNN----- 143
QY      242  GTGISPTVNDNPEKRLVKEIRICEVRRPGCPVYSSLKGGKCSKTKKSPPEVRFYAGC 301
Db      144  GLGVSTRISNVNACWPEQESRLCNLRPCVDYDITLILKAGKKCLAVYQPEASMTLAC 203
QY      302  SSVKKYRPKYCGSCVDGRCTPLQTRTVKMRFRCEDEGEMFSKNVMTIOSCKCNYNCPHPN 361
Db      204  ISTRSYQPKYCGVCMNRCCIPYKSKTIDVSFQCPDGLGFSROYLWINACFCNLSCRPN 263
QY      362  EASFRLXSLFNDIHKFRD 379
Db      264  D-----IFADLESYPD 274
```

Search completed: August 5, 2003, 14:06:50
Job time : 35.9132 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:02:12 ; Search time 39.1026 Seconds
(without alignments)
1546.567 Million cell updates/sec

Title: US-09-495-448a-4
Perfect score: 2116
Sequence: 1 MSSRIARALALVITLHLTR.....ANEAPFPYRLFNDHKFRD 381

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_19Jun03.*
2: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2116	100.0	381	18	AAW35730
2	2116	100.0	381	22	AAE05921
3	2106	99.5	381	18	AAW35957
4	2106	99.5	381	22	AAE05921
5	2106	99.5	381	23	AAU79761
6	2106	99.5	381	23	AAE18107
7	2106	99.5	381	23	ABO5438
8	2106	99.5	455	21	AAE43987
9	2098	99.1	381	23	ABG76937

10	1971.5	93.2	374	23	ABO9202	HCCE CNN family pr
11	1971.5	93.2	375	17	AAE0919	Connective tissue
12	1971.5	93.2	375	20	AAE31620	Human CTGF-2. Hom
13	1971.5	93.2	375	23	AAE18108	Human alternative
14	1929	91.2	379	13	AAE2565	Beta-IG-M1. Mus m
15	1929	91.2	379	22	AAE05920	Mouse cytochrome-c
16	1929	91.2	379	23	ABO9201	Mouse cytochrome-c
17	1695	80.1	375	22	AAE05939	Chicken CEP10 prot
18	1695	80.1	375	23	ABO9203	Chicken ce10 CNN f
19	953	45.0	347	20	AAE24379	Rat connective tis
20	952.5	45.0	348	13	AAE25566	Beta-IG-M2. Mus m
21	948.5	44.8	348	18	AAW35731	Murine Fisp12. Mu
22	948.5	44.8	348	21	AAE44756	Mouse connective t
23	948.5	44.8	348	22	AAE05922	Mouse fibroblast s
24	948.5	44.8	348	23	ABO9205	Fisp-12 CNN family
25	948	44.8	347	18	AAW12694	Connective tissue
26	948	44.8	347	21	AAE33340	Amino acid sequenc
27	947	44.8	349	16	AAE79964	Connective tissue
28	947	44.8	349	18	AAE09089	Human connective t
29	947	44.8	349	18	AAW11302	Connective tissue
30	947	44.8	349	19	AAE62084	Human connective t
31	947	44.8	349	20	AAE18361	Human connective t
32	947	44.8	349	20	AAW81425	Connective tissue
33	947	44.8	349	21	AAE29339	Human connective t
34	947	44.8	349	21	AAE29340	Human connective t
35	947	44.8	349	21	AAE44755	Human connective t
36	947	44.8	349	22	AAE05923	Human connective t
37	947	44.8	349	22	AAE4598	Amino acid sequenc
38	947	44.8	349	22	AAE0791	Human shear stress
39	947	44.8	349	22	AAE0664	Human connective t
40	947	44.8	349	22	AAE8831	Human connective t
41	947	44.8	349	23	ABP68624	Human pancreatic c
42	938.5	44.4	348	23	ABO9204	Human ctyf CNN fam
43	869	40.2	349	21	AAE44754	Bovine connective
44	851.5	40.2	351	14	AAE31599	Chicken nov protei
45	851.5	40.2	351	22	AAE05940	Avian Nov protein.

ALIGNMENTS

RESULT 1	AAW35730	standard; Protein: 381 AA.
ID	AAW35730	
AC	AAW35730;	
XX		
DT	27-MAR-1998	(first entry)
XX		
DE	Human cysteine rich protein 61 (Cyr61).	
XX		
KW	Cysteine rich protein 61; Cyr61; human;	
KW	extracellular matrix signalling molecule; cell adhesion;	
KW	cell migration; cell proliferation; angiogenesis; chondrogenesis;	
KW	oncogenesis; haematostasis; wound healing; organ regeneration.	
XX		
OS	Homo sapiens.	
XX		
PN	W09733995-A2.	
XX		
PD	18-SEP-1997.	
XX		
PE	14-MAR-1997;	97WO-US04193.
XX		
PR	15-MAR-1996;	96US-0013958.
XX		
PA	(MUNI-) MUNIN CORP.	
XX		
PI	Lau LF;	
XX		
DR	WPI: 1997-470875/43.	
XX	N-PSDB: AAT94699.	

PT Isolated and purified cysteine rich protein 61, Cyr61 - useful to
 PT modulate e.g. haemostasis, induce wound healing, promote organ
 PT regeneration etc
 XX
 PS Claim 2; Page 112-113; 133pp; English.
 XX
 CC This protein sequence comprises human cysteine rich protein 61
 CC (Cyr61), an extracellular matrix signalling molecule. Its amino
 CC acid sequence was deduced from a human placental cDNA clone (see
 CC AAT94699). Cyr61 polypeptides can be expressed in transformed or
 CC transfected host cells. Cyr61 can be used to modulate
 CC haemostasis, induce wound healing in a tissue, promote organ
 CC regeneration, improve tissue grafting or promote bone or prothesis
 CC implantation (claimed). It can also be used to screen for a
 CC modulator of angiogenesis, chondrogenesis, oncogenesis, cell
 CC adhesion, cell migration, cell proliferation, expand a population
 CC of undifferentiated haematopoietic stem cells in culture and to
 CC screen for a mitogen (claimed). Ex vivo methods for using
 CC mammalian extracellular matrix signalling molecules to prepare
 CC blood products are also provided.
 XX
 XX Sequence 381 AA;

Query Match 100.0%; Score 2116; DB 18; Length 381;
 Best Local Similarity 100.0%; Pred. No. 1.9e-162;
 Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSRIARALALVYTLHLTRALSTCPACCHPLEAPKCAPGVGLVNDGCGCKVCANOL 60
 DB 1 MSSRIARALALVYTLHLTRALSTCPACCHPLEAPKCAPGVGLVNDGCGCKVCANOL 60
 QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPEYNSRIYQNGESFQPCNQHQ 120
 DB 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPEYNSRIYQNGESFQPCNQHQ 120
 QY 121 CTCIDAVGCIPLCPQELSLPNIGCPNPLVKTGGCCCEWVCDSDSKPMEDDGLG 180
 DB 121 CTCIDAVGCIPLCPQELSLPNIGCPNPLVKTGGCCCEWVCDSDSKPMEDDGLG 180
 QY 181 KELGFDASEVELTRNNELIANGKRSRLRPFVGMERILYNPLOGOKCIYQTTSMGQS 240
 DB 181 KELGFDASEVELTRNNELIANGKRSRLRPFVGMERILYNPLOGOKCIYQTTSMGQS 240
 QY 241 KTCGTGISTRYTNDNPECLVKEIRICEVRPGQPYVSSILKKKCKSKTKKSPVAFYTY 300
 DB 241 KTCGTGISTRYTNDNPECLVKEIRICEVRPGQPYVSSILKKKCKSKTKKSPVAFYTY 300
 QY 301 AGCLSVKRYKPKYCGSCVDRCCTPOLTRTVKMRFRCEDEFTSKNVMIOGCKNVCNP 360
 DB 301 AGCLSVKRYKPKYCGSCVDRCCTPOLTRTVKMRFRCEDEFTSKNVMIOGCKNVCNP 360
 QY 361 HANEAAPEFYRLFNDIHKFRD 381
 DB 361 HANEAAPEFYRLFNDIHKFRD 381

RESULT 2
 AAE05921
 ID AAE05921 standard; Protein; 381 AA.
 XX
 AC AAE05921;

XX 24-SEP-2001 (first entry)
 DE Human cysteine-rich protein (Cyr61).
 XX
 KM Cysteine-rich protein; Cyr61; extracellular matrix signalling molecule;
 KM fibroblast secreted protein; Fisp12; connective tissue growth factor;
 KM CTGF; ECM; cell adhesion; cell migration; fibroblast cell proliferation;
 KM angiogenesis; wound healing; integrin receptor; atherosclerosis; tumour;
 KM heart disease; fibrosis; gene therapy; human.
 XX
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH Region 163..229
 FT /note="Cysteine free region"
 FT Domain 212..281
 FT /note="Domain III"
 FT Domain 282..381
 FT /note="Domain IV"
 XX
 PN W0200155210-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 31-JAN-2001; 2001WO-US03267.
 XX
 PR 31-JAN-2000; 2000US-0495448.
 PR 15-MAY-2000; 2000US-0204364.
 PR 06-OCT-2000; 2000US-0238705.
 XX
 PA (MUNN-) MUNN CORP.
 XX
 PI Lau LF, Yeung C, Greenspan JA;
 XX
 DR WPI; 2001-465561/50.
 DR N-PDB; AAD11221.
 XX

PT Novel human cysteine-rich protein 61 (Cyr61) fragment useful in methods
 PT for screening for modulators of cell adhesion, fibroblast cell
 PT proliferation, angiogenesis and cell migration
 XX
 PS Claim 30; Page 171-172; 186pp; English.

XX The invention relates to extracellular matrix (ECM) signalling
 CC molecules involved in cellular response to growth factors. More
 CC particularly the invention is directed to cysteine-rich protein
 CC (Cyr61), and Cyr61-related proteins such as fibroblast secreted
 CC protein (Fisp12) and connective tissue growth factor (CTGF) and
 CC nucleic acid molecules encoding such proteins. The polypeptides
 CC of the invention are members of cysteine-rich secreted protein
 CC family. Human Cyr61 fragment is useful in methods for screening
 CC modulators of cell adhesion, cell migration, fibroblast cell
 CC proliferation, angiogenesis, wound healing and Cyr61-integrin
 CC receptor interaction. Modulator of Cyr61-integrin alphavbeta3
 CC interaction is used for the preparation of a medicament for the
 CC treatment of atherosclerosis, heart disease, tumour metastasis,
 CC fibrosis, tumour growth, disorders associated with inadequate
 CC angiogenesis; aberrant granulation tissue development; aberrant
 CC fibroblast growth and wounds. Polynucleotides of the invention
 CC are useful in gene therapy. The present sequence is human Cyr61
 CC protein.
 XX
 XX Sequence 381 AA;

Query Match 100.0%; Score 2116; DB 22; Length 381;
 Best Local Similarity 100.0%; Pred. No. 1.9e-162;
 Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSRIARALALVYTLHLTRALSTCPACCHPLEAPKCAPGVGLVNDGCGCKVCANOL 60
 DB 1 MSSRIARALALVYTLHLTRALSTCPACCHPLEAPKCAPGVGLVNDGCGCKVCANOL 60
 QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPEYNSRIYQNGESFQPCNQHQ 120
 DB 61 NEDCSKTQPCDHTKGLGECNFGASSTALKGICRAQSEGRPEYNSRIYQNGESFQPCNQHQ 120
 QY 121 CTCIDAVGCIPLCPQELSLPNIGCPNPLVKTGGCCCEWVCDSDSKPMEDDGLG 180
 DB 121 CTCIDAVGCIPLCPQELSLPNIGCPNPLVKTGGCCCEWVCDSDSKPMEDDGLG 180
 QY 181 KELGFDASEVELTRNNELIANGKRSRLRPFVGMERILYNPLOGOKCIYQTTSMGQS 240
 DB 181 KELGFDASEVELTRNNELIANGKRSRLRPFVGMERILYNPLOGOKCIYQTTSMGQS 240

QY 241 KTCGTGISTRYTNDNPECLVKEIRICEVRCGQPVYSLKGGKCKSTKKSPPEVPTTY 300
 DB 241 KTCGTGISTRYTNDNPECLVKEIRICEVRCGQPVYSLKGGKCKSTKKSPPEVPTTY 300
 QY 301 AGCLSVKKYRPRKYGSCVDGRCCTPQLTRTVYKMRFCEDGETFSKNVMMIOSCKKNYCP 360
 DB 301 AGCLSVKKYRPRKYGSCVDGRCCTPQLTRTVYKMRFCEDGETFSKNVMMIOSCKKNYCP 360
 QY 361 HANEAFPEYRLFNDIHKFRD 381
 DB 361 HANEAFPEYRLFNDIHKFRD 381

RESULT 3

AAM35957
 ID AAM35957 standard; Protein: 381 AA.

AAW35957;

DT 05-MAR-1998 (first entry)

DE Human monocyte mature differentiation factor.

KW Human; monocyte; mature; differentiation factor; MMD; macrophage;
 cancer; immune activator; tissue culture; infectious disease.

OS Homo sapiens.

PN JP09234079-A.

PD 09-SEP-1997.

PE 04-MAR-1996; 96JP-0075236.

PR 04-MAR-1996; 96JP-0075236.

PA (TOYM) TOYOBO KK.

DR WPI: 1997-497320/46.

DR N-PSDB; AAT97142.

PT A monocyte mature differentiation factor - useful for the long term

PS tissue culture of macrophage(s)

PS Claim 9; Page 12-13; 22pp; Japanese.

CC The present sequence represents a monocyte mature differentiation

CC factor (MMD) which maintains the life of macrophages for long periods

CC in liquid culture. MMD can be used as an anti-cancer agent, an immune

CC activator and to treat infectious diseases.

SO Sequence 381 AA;

Query Match 99.5%; Score 2106; DB 18; Length 381;

Best Local Similarity 99.5%; Pred. No. 1,2e-161; Mismatches 1; Indels 0; Gaps 0;

Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALALVTLHLTRLALSTCPAACHCPLFAKCAAGVGLVRDGGCCCKVCAKOL 60
 DB 1 MSSRIARALALVTLHLTRLALSTCPAACHCPLFAKCAAGVGLVRDGGCCCKVCAKOL 60
 QY 61 NEDSKTOPCDHTKGLGCNFGASSTALKGICRAOSEGRPCENSRITONGSEFPQNCQH 120
 DB 61 NEDSKTOPCDHTKGLGCNFGASSTALKGICRAOSEGRPCENSRITONGSEFPQNCQH 120
 QY 121 CTCIDGAVGCIPLCPOELSLPNLGCNPRILVYKVGCCCEWVDCEDSIRKPMEDQDLG 180
 DB 121 CTCIDGAVGCIPLCPOELSLPNLGCNPRILVYKVGCCCEWVDCEDSIRKPMEDQDLG 180
 QY 181 KETGFDASEVELTRNNELIANGKRSILKRLPVFGMEBRILYNPLQGCKCIYQTTSMGSCS 240
 DB 181 KETGFDASEVELTRNNELIANGKRSILKRLPVFGMEBRILYNPLQGCKCIYQTTSMGSCS 240

QY 241 KTCGTGISTRYTNDNPECLVKEIRICEVRCGQPVYSLKGGKCKSTKKSPPEVPTTY 300
 DB 241 KTCGTGISTRYTNDNPECLVKEIRICEVRCGQPVYSLKGGKCKSTKKSPPEVPTTY 300
 QY 301 AGCLSVKKYRPRKYGSCVDGRCCTPQLTRTVYKMRFCEDGETFSKNVMMIOSCKKNYCP 360
 DB 301 AGCLSVKKYRPRKYGSCVDGRCCTPQLTRTVYKMRFCEDGETFSKNVMMIOSCKKNYCP 360
 QY 361 HANEAFPEYRLFNDIHKFRD 381
 DB 361 HANEAFPEYRLFNDIHKFRD 381

RESULT 4

AAB90773
 ID AAB90773 standard; Protein: 381 AA.

AAB90773;

DT 15-JUN-2001 (first entry)

DE Human shear stress-response protein SEQ ID NO: 46.

KW Human; shear stress-response protein; vascular disease;
 arteriosclerosis.

OS Homo sapiens.

PN WO200125427-A1.

PD 12-APR-2001.

PE 02-OCT-2000; 2000WO-JP06840.

PR 01-OCT-1999; 99JP-0280976.

PA (KYOW) KYOWA HAKKO KOGYO KK.

PA (NOJI/) NOJIMA H.

PI Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;

PI Kuga T, Sekine S, Nakamura Y, Sugano S;

DR WPI: 2001-266308/27.

DR N-PSDB; AAH02896.

PT DNA sequences, proteins encoded by them and antibodies against them

PT useful in diagnosis and treatment of vascular disease caused by

PS arteriosclerosis -

PS Claim 60; Page 345-346; 678pp; Japanese.

CC The present invention provides the protein and coding sequences of a

CC number of human shear stress response proteins. These are useful in the

CC diagnosis, treatment and screening of vascular diseases caused by

CC arteriosclerosis, including heart failure, post-PTCA restenosis and

CC hypertension.

SO Sequence 381 AA;

Query Match 99.5%; Score 2106; DB 22; Length 381;

Best Local Similarity 99.5%; Pred. No. 1,2e-161; Mismatches 1; Indels 0; Gaps 0;

Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALALVTLHLTRLALSTCPAACHCPLFAKCAAGVGLVRDGGCCCKVCAKOL 60
 DB 1 MSSRIARALALVTLHLTRLALSTCPAACHCPLFAKCAAGVGLVRDGGCCCKVCAKOL 60
 QY 61 NEDSKTOPCDHTKGLGCNFGASSTALKGICRAOSEGRPCENSRITONGSEFPQNCQH 120
 DB 61 NEDSKTOPCDHTKGLGCNFGASSTALKGICRAOSEGRPCENSRITONGSEFPQNCQH 120
 QY 121 CTCIDGAVGCIPLCPOELSLPNLGCNPRILVYKVGCCCEWVDCEDSIRKPMEDQDLG 180
 DB 121 CTCIDGAVGCIPLCPOELSLPNLGCNPRILVYKVGCCCEWVDCEDSIRKPMEDQDLG 180

Db	12.	CTCICIDGAVGCIPLCPQELSLPNLGCNPRLVAVTGGCCCEMWCCDEBSIKDPMEDQGLG	180
QY	181	KELGFPDASEVELTRNNELIAVGKSGSLKRLPLFGHEPRLILNPLDQKCIYOTTSMQCS	240
Db	181	KELGFPDASEVELTRNNELIAVGKSGSLKRLPLFGHEPRLILNPLDQKCIYOTTSMQCS	240
QY	241	KTCGTGISTPVTINDNPECLRYKETRICEVRPCGQVYSSLLKGGKCSSTKKSPEPVFTY	300
Db	241	KTCGTGISTPVTINDNPECLRYKETRICEVRPCGQVYSSLLKGGKCSSTKKSPEPVFTY	300
QY	301	AGCLSVKKYKPKYCGSCVDRCCTPOLRTYKMFRCDEGETFSKNVMIQSCKNYCP	360
Db	301	AGCLSVKKYKPKYCGSCVDRCCTPOLRTYKMFRCDEGETFSKNVMIQSCKNYCP	360
QY	361	HANEAAPEFRLFNDIHKFRD 381	
Db	361	HANEAAPEFRLFNDIHKFRD 381	

XX	AAU79761 standard; Protein; 381 AA.
XX	
XX	AAU79761;
XX	
XX	30-JUL-2002 (first entry)
XX	
XX	Human Cyt61 protein.
XX	
XX	Human; uterine leiomyoma proliferation; uterine leiomyoma formation;
XX	Cyt61; cytosolic.
XX	
XX	Homo sapiens.
XX	
XX	WO200226193-A2.
XX	
XX	04-APR-2002.
XX	
XX	28-SEP-2001; 2001MO-US30783.
XX	
XX	29-SEP-2000; 2000US-236887P.
XX	
XX	(AMHP) AMERICAN HOME PROD CORP.
XX	
XX	Zhang Z, Sampath D, Zhu Y, Winneker R;
XX	
XX	WPI; 2002-383245/A1.
XX	
XX	N-PSDB; ABR48899.
XX	
XX	Preventing uterine leiomyoma formation or inhibiting proliferation of
XX	uterine leiomyoma in subject, comprises modulating or increasing the
XX	level of Cyt61 in leiomyoma tissue
XX	
XX	Disclosure; Fig 6; 92pp; English.
XX	
XX	The present invention relates to a method of inhibiting proliferation
XX	of uterine leiomyoma or preventing uterine leiomyoma formation. The
XX	method comprises increasing the level of Cyt61 in leiomyoma tissue.
XX	The invention also describes compounds and compositions that stimulate
XX	induction of the Cyt61 gene and compounds that increase Cyt61 activity.
XX	The compositions and the method of the invention are useful for
XX	preventing uterine leiomyoma formation or inhibiting proliferation of
XX	uterine leiomyoma in a subject. The method is particularly useful for
XX	treating or preventing uterine leiomyoma formation, or inhibiting
XX	proliferation of uterine leiomyoma in a subject. The present sequence
XX	represents human Cyt61.
XX	Note: The present sequence shown in Fig 6 is not shown in the correct
XX	sequence order in the figure. The start of the sequence is shown on
XX	page 8/10 of the figures and the rest of the sequence is shown on
XX	page 6/10 of the figures.
XX	
XX	Sequence 381 AA;
XX	

Query Match	Similarity	99.5%	Score	2106;	DB	23;	Length	381;
Best Local	Similarity	99.5%	Pred. No.	1,2e-161;				
Matches	379;	Conservative	1;	Mismatches	1;	Indels	0;	Gaps
QY	1	MSRIRALALVTLHLTRLALSTCPAACHCPLCAPGVGLVROGCGCCKVCACAKOL	60					
DB	1	MSRIRALALVTLHLTRLALSTCPAACHCPLCAPGVGLVROGCGCCKVCACAKOL	60					
QY	61	NEDCSKTOPCHTKLECNFGASSALGICICANOSBGRCEFNSRIYONGSEFOPMCKHQ	120					
DB	61	NEDCSKTOPCHTKLECNFGASSALGICICANOSBGRCEFNSRIYONGSEFOPMCKHQ	120					
QY	121	CTCIDGAVGCIPLCPQBELSLPNLGCGPNPLVYKVGQCCBWCDEDSIKDPMEDDGLLG	180					
DB	121	CTCIDGAVGCIPLCPQBELSLPNLGCGPNPLVYKVGQCCBWCDEDSIKDPMEDDGLLG	180					
QY	181	KELGDAEVELTRNNELLAVGKGGSKRLPVFGMEPRLLYNPLDGOCKIYOTTSWQOS	240					
DB	181	KELGDAEVELTRNNELLAVGKGGSKRLPVFGMEPRLLYNPLDGOCKIYOTTSWQOS	240					
QY	241	KTCGGISTRYTNDNDECRILVETRICERPCQGVYSILKKGKCSKTKKSPDEPRFTY	300					
DB	241	KTCGGISTRYTNDNDECRILVETRICERPCQGVYSILKKGKCSKTKKSPDEPRFTY	300					
QY	301	AGCLSVKKYRPKYCGSCVDGRCTPOLTRTYAMRFCEDEGETFSKNVMMIIOGCKNYPNC	360					
DB	301	AGCLSVKKYRPKYCGSCVDGRCTPOLTRTYAMRFCEDEGETFSKNVMMIIOGCKNYPNC	360					
QY	361	HANEAPFYRLFNDIHKFRD 381						
DB	361	HANEAPFYRLFNDIHKFRD 381						

ID	AAE18107 standard; Protein; 361 AA.
XX	AAE18107;
DT	07-MAY-2002 (first entry)
DE	Human connective tissue growth factor-2 (CTGF-2).
XX	
KW	Human; angiogenesis; connective tissue growth factor-2; CTGF-2; tumour;
KW	ischemia; restenosis; tissue repair; wound healing; congenital defect;
KW	cardiovascular disease; atherosclerosis; heart failure; angina; trauma;
KW	burns; osteoporosis; periodontal disease; liver failure; tranquilizer;
KV	vulnerary; cosmetic plastic surgery; vasotrophic; hepatotropic; ulcer;
gene therapy.	
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Region
FT	/note= "Immunogenic epitope"
FT	47..51
FT	Region
FT	/note= "Immunogenic epitope"
FT	59..75
FT	Region
FT	/note= "Immunogenic epitope"
FT	91..119
FT	Region
FT	/note= "Immunogenic epitope"
FT	145..150
FT	Region
FT	/note= "Immunogenic epitope"
FT	164..176
FT	Region
FT	/note= "Immunogenic epitope"
FT	202..208
FT	Region
FT	/note= "Immunogenic epitope"
FT	223..228
FT	Region
FT	/note= "Immunogenic epitope"
FT	239..244
FT	Region
FT	/note= "Immunogenic epitope"
FT	250..257
FT	Region
FT	/note= "Immunogenic epitope"

FT Region 279..296
 FT /note= "Immunogenic epitope"
 FT Region 307..314
 FT /note= "Immunogenic epitope"
 FT Region 318..323
 FT /note= "Immunogenic epitope"
 FT Region 337..343
 FT /note= "Immunogenic epitope"
 XX
 PN WO200204480-A2.
 XX
 PD 17-JAN-2002.
 XX
 PE 11-JUL-2001; 2001WO-US21799.
 XX
 PR 11-JUL-2000; 2000US-217402P.
 PR 18-MAY-2001; 2001US-291642P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (TRGE) TRANSGENE SA.
 XX
 PI Li H, Adams MD, Calenda V, Fataccioli V;
 DR WPI; 2002-1171698/22.
 DR N-PDB; AAD29095.
 XX
 PT Stimulating angiogenesis in a mammal preferably human having ischemia
 PT or restenosis or is treated for limb revascularization, by
 PT administering connective tissue growth factor-2 polypeptide or
 PT polynucleotide
 XX
 PS Example 1; Fig 1; 131pp; English.
 XX
 CC The present invention relates to a method for stimulating angiogenesis in
 CC a mammal. The method comprises administering a polynucleotide encoding
 CC connective tissue growth factor-2 (CTGF-2) or an active fragment or its
 CC derivative. The method is useful for stimulating angiogenesis in a mammal
 CC preferably human having ischemia or restenosis or is treated for limb
 CC revascularization which is leg or arm. The invention is useful for
 CC inhibiting tumor growth, where angiogenesis is utilised for enhancing
 CC the repair of connective and support tissue, promoting the attachment,
 CC fixation and stabilisation of tissue implants and enhancing wound
 CC healing, hence is useful for treating cardiovascular disease e.g.
 CC atherosclerosis, reperfusion injury such as heart failure, angina,
 CC ischemia; and is also used to differentiate, proliferate and attract
 CC cells leading to regeneration of tissues which is utilised to repair
 CC replace or protect tissue damaged by congenital defects, trauma (burns,
 CC ulcer, etc), age, disease (e.g. osteoporosis, periodontal disease,
 CC liver failure), surgery including cosmetic plastic surgery. The present
 CC sequence is human CTGF-2. CTGF-2 gene is useful in gene therapy.
 XX
 SQ Sequence 381 AA;
 Query Match 99.5%; Score 2106; DB 23; Length 381;
 Best Local Similarity 99.5%; Pred. No. 1.2e-161;
 Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 241 KTCGTGISTRVNDNPECRLVETRICVPRGCPVYSLLKGGKCKSKTKSPVPVPTTY 300
 DB 241 KTCGTGISTRVNDNPECRLVETRICVPRGCPVYSLLKGGKCKSKTKSPVPVPTTY 300
 QY 301 AGCLSVKRYKRYKCGSCVDRCCTPOLTRVKNMFRCEDEDETSKNVMYIOSCKYNNCP 360
 DB 301 AGCLSVKRYKRYKCGSCVDRCCTPOLTRVKNMFRCEDEDETSKNVMYIOSCKYNNCP 360
 QY 361 HANEAAPFYRLFNDRHFRD 381
 DB 361 HANEAAPFYRLFNDRHFRD 381
 RESULT 7
 ABB05438
 ID ABB05438 standard; Protein; 381 AA.
 XX
 AC ABB05438;
 XX
 DT 15-APR-2002 (first entry)
 XX
 DE Human Cyr61 protein SEQ ID NO:2.
 XX
 KW Human; Cyr61; breast cancer; sex steroid receptor; cytosolic; promoter;
 KW sex steroid response element; cysteine rich heparin-binding protein;
 KW cell proliferation; heparin binding epidermal growth factor;
 KW epidermal growth factor; basic fibroblastic growth factor.
 XX
 OS Homo sapiens.
 XX
 PN WO200198359-A2.
 PD 27-DEC-2001.
 XX
 PE 21-JUN-2001; 2001WO-US19823.
 XX
 PR 21-JUN-2000; 2000US-213182P.
 PR 16-MAY-2001; 2001US-291510P.
 XX
 PA (AMHP) AMERICAN HOME PROD CORP.
 XX
 PI Sampath D, Zhang Z, Winneker R;
 DR WPI; 2002-147796/19.
 DR N-PDB; ABA93127, ABA93130.
 XX
 PT Regulation of Cyr61 expression and activity for preventing and
 PT inhibiting breast cancer comprises use of a Cyr61 neutralizing
 PT antibody, an anti-sense oligonucleotide and an antibody which
 XX
 PS Claim 6; Fig 1; 86pp; English.
 XX
 CC The present invention describes a method for the prevention or inhibition
 CC of breast cancer cell proliferation. The method comprises administration
 CC of a compound that inhibits the interaction of a sex steroid receptor
 CC with a sex steroid response element of the Cyr61 (cysteine rich heparin-
 CC binding protein) promoter. Cyr61 has cytosolic activity. An antibody (1)
 CC which neutralises Cyr61 can be used to prevent or inhibit breast cancer
 CC cell proliferation by blocking sex steroid induced and growth factor
 CC induced synthesis of Cyr61 DNA, where the growth factor is epidermal,
 CC heparin binding epidermal or basic fibroblastic growth factor. (1) can be
 CC used to diagnose or stage breast cancer where the level of Cyr61 in a
 CC positive/suspect breast cancer cell is compared to the level in a
 CC normal cell; an increase in the level of Cyr61 compared to the level in
 CC normal tissue indicates the presence of breast cancer. The level of
 CC Cyr61 being determined by exposing the tissues to (1), and an increase
 CC in the level of bound antibody by the suspect/positive cell as compared
 CC to the normal tissue indicates the presence of breast cancer. The present
 CC sequence represents the human Cyr61 protein, which is used in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 381 AA;

Query Match 99.5%; Score 2106; DB 23; Length 381;
Best Local Similarity 99.5%; Pred. No. 1.2e-161;
Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALAVVTLHLTRLALSTCPAACHCPLPAPKAPGVGLVRDGGCCCKVCAQOL 60
DB 1 MSSRIARALAVVTLHLTRLALSTCPAACHCPLPAPKAPGVGLVRDGGCCCKVCAQOL 60
QY 61 NEDCSKTQPCDHTKLECNFGASSTALNGICRAOSEGRPCENSRITYNGESFQPNCHQ 120
DB 61 NEDCSKTQPCDHTKLECNFGASSTALNGICRAOSEGRPCENSRITYNGESFQPNCHQ 120
QY 121 CTCIDGAVGICPLCPQELSLPMLGCPNPLVYVGTGCCCEWYCDSDSIKDPMEDDGLLG 180
DB 121 CTCIDGAVGICPLCPQELSLPMLGCPNPLVYVGTGCCCEWYCDSDSIKDPMEDDGLLG 180
QY 181 KELGFDASEVELTRNNELIANGKGRSLKRLPVFGMEPRILYNPLDGGKCIYOTTSWSQCS 240
DB 181 KELGFDASEVELTRNNELIANGKGRSLKRLPVFGMEPRILYNPLDGGKCIYOTTSWSQCS 240
QY 241 KTCGIGISTRVTNDNPECRIVKETRICEVRPCGQPVYSSLKKGKCSKTKSPPEVRYTY 300
DB 241 KTCGIGISTRVTNDNPECRIVKETRICEVRPCGQPVYSSLKKGKCSKTKSPPEVRYTY 300
QY 301 AGCLSVKKRYKRYCGSCVDGRCTPOLTRTYVMRRFCEDGETFSKNVMIOSCKNYNCP 360
DB 301 AGCLSVKKRYKRYCGSCVDGRCTPOLTRTYVMRRFCEDGETFSKNVMIOSCKNYNCP 360
QY 361 HANEAAFPYRLFNDIHKFRD 381
DB 361 HANEAAFPYRLFNDIHKFRD 381

RESULT 8
AAB43987
ID AAB43987 standard; Protein: 455 AA.
XX AAB43987;
DT 08-FEB-2001 (first entry)
DE Human cancer associated protein sequence SEQ ID NO:1432.
XX

Human; cancer associated gene; cancer antigen; detection; cancer;
diagnosis; cytostatic; proliferative; vulnerrary; immunomodulator;
antidiabetic; antiaesthetic; antineuritic; antitumor; antiviral;
dermatological; neuroprotective; antiallergic; antibacterial; cardiac;
vasotrophic; antipsoriatic; antidiabetic; gene therapy; inflammation;
immune disorder; haematopoietic cell disorder; autoimmune disorder;
allergic reaction; graft versus host disease; organ rejection;
neurological disease; drug screening.

OS Homo sapiens.
XX
XX WO200055350-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US05882.
XX
XX 12-MAR-1999; 99US-0124270.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Ruben SM;
XX
XX WPI: 2000-587533/55.
XX
XX DR N-PSDB: AAC78196.
XX
XX PT Novel isolated nucleic acids comprising sequences encoding peptides
useful for treating or diagnosing e.g. cancer -

XX Claim 11; Page 2116-2118; 2352pp; English.
PS AAC7607 to AAC78448 encode the human cancer associated proteins given
XX in AAB4398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnerrary; immunomodulator;
CC antidiabetic; antiaesthetic; antineuritic; antitumor; antiviral;
CC antiinflammatory; antiallergic; antibacterial; cardiac;
CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;
CC vasotrophic; antipsoriatic and antidiabetic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC7849 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention.

Sequence 455 AA:
SQ

Query Match 99.5%; Score 2106; DB 21; Length 455;
Best Local Similarity 99.5%; Pred. No. 1.5e-161;
Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALAVVTLHLTRLALSTCPAACHCPLPAPKAPGVGLVRDGGCCCKVCAQOL 60
DB 75 MSSRIARALAVVTLHLTRLALSTCPAACHCPLPAPKAPGVGLVRDGGCCCKVCAQOL 134
QY 61 NEDCSKTQPCDHTKLECNFGASSTALNGICRAOSEGRPCENSRITYNGESFQPNCHQ 120
DB 135 NEDCSKTQPCDHTKLECNFGASSTALNGICRAOSEGRPCENSRITYNGESFQPNCHQ 194
QY 121 CTCIDGAVGICPLCPQELSLPMLGCPNPLVYVGTGCCCEWYCDSDSIKDPMEDDGLLG 180
DB 195 CTCIDGAVGICPLCPQELSLPMLGCPNPLVYVGTGCCCEWYCDSDSIKDPMEDDGLLG 254
QY 241 KTCGIGISTRVTNDNPECRIVKETRICEVRPCGQPVYSSLKKGKCSKTKSPPEVRYTY 300
DB 315 KTCGIGISTRVTNDNPECRIVKETRICEVRPCGQPVYSSLKKGKCSKTKSPPEVRYTY 374
QY 301 AGCLSVKKRYKRYCGSCVDGRCTPOLTRTYVMRRFCEDGETFSKNVMIOSCKNYNCP 360
DB 375 AGCLSVKKRYKRYCGSCVDGRCTPOLTRTYVMRRFCEDGETFSKNVMIOSCKNYNCP 434
QY 361 HANEAAFPYRLFNDIHKFRD 381
DB 435 HANEAAFPYRLFNDIHKFRD 455

RESULT 9
ABG76937
ID ABG76937 standard; Protein: 381 AA.
XX ABG76937;
XX
XX 05-NOV-2002 (first entry)
XX
XX Human protein, comprising CYR61, designated SECL.
XX
XX
XX Human; SEC; NOV; immunosuppressive; hepatotropic;
XX antiinflammatory; angiogenic-associated disorder; diagnostic;
XX gene therapy; developmental disorder; immune disease;
KW

KW signal transduction pathway disorder; metabolic disorder;
 KW feeding disorder; obesity; wasting disorder; neurodegenerative disorder;
 KW Alzheimer's disease; Parkinson's disease; behavioural disorder; allergy;
 KW asthma; atherosclerosis; cardiomyopathy; angina pectoris;
 KW autoimmune disease; retinal disease; cirrhosis; diabetes;
 KW infectious disease; human immunodeficiency virus; HIV; cancer;
 KW hyperostosis; hypoplasia; multiple sclerosis; urinary retention;
 KW osteoporosis; Crohn's disease; ulcer; neurological disorder; anxiety;
 KW haemophilia; cirrhosis; immunogen; vaccine.
 XX Homo sapiens.
 OS
 XX MO20025705-A2.
 PN
 XX 18-JUL-2002.
 PD
 XX 11-JAN-2002; 2002WO-US00609.
 PF
 XX 11-JAN-2001; 2001US-261013P.
 PR 11-JAN-2001; 2001US-261014P.
 PR 11-JAN-2001; 2001US-261018P.
 PR 11-JAN-2001; 2001US-261026P.
 PR 11-JAN-2001; 2001US-261029P.
 PR 17-AUG-2001; 2001US-313170P.
 PR 10-SEP-2001; 2001US-318410P.
 XX
 XX (CURA-) CURAGEN CORP.
 PA
 PI Mezes PS, Rastell L, Hermann JL, MacDougall JR, Zhong H;
 PI Casman SJ, Boldo F, Shmukets RA, Gorman L, Crasta OR, Mysore KK;
 PI Folkerts O, Martin GB, Eisen A, Spaderna SK, Vermet CAM, Bergh C;
 PI Spletter KA, DiPippo VA, Zernusen BD, Peyman JA, Ellemann K;
 PI Stone DJ, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK;
 PI Burgess CE, Edinger S.
 XX
 XX WPI: 2002-550675/63.
 DR N-PSDB; ABS59522.
 DR
 XX Human SECX/NOVX polypeptide useful for diagnosing, preventing or
 PT treating disorders associated with aberrant expression or activity of
 PT SECX/NOVX nucleic acids and proteins e.g., diabetes
 PT
 XX
 XX Example 3; Page 9; 443pp; English.
 XX
 CC The invention discloses the isolated human polypeptides, and
 CC polynucleotides encoding them, that have been designated SECX and NOVX.
 CC The polypeptides can be used for treating, or delaying, the onset of an
 CC antigenic-associated disorder or treating a pathological state in a
 CC subject, preferably a mammal. They can also be used in determining the
 CC presence of, or predisposition to, a disease associated with altered
 CC levels of the polypeptides and polynucleotides of any one of the 12
 CC sequences (SEC1-12), for raising antibodies, for identifying an agent
 CC that binds to, or that modulates the expression or activity of the
 CC polypeptide, for treating or preventing a NOVX-associated disorder
 CC (NOV1-8) and as a pharmaceutical composition comprising the polypeptide,
 CC polynucleotide or the antibody. The polypeptides and polynucleotides are
 CC useful in diagnostic applications where their amounts are assessed, or
 CC for the manufacture of a medicament (e.g. gene therapy) for treating or
 CC preventing disorders or syndromes such as developmental disorders, immune
 CC diseases, signal transduction pathway disorders, metabolic disorders,
 CC feeding disorders (including obesity), wasting disorders,
 CC neurodegenerative disorders (including Alzheimer's disease and
 CC Parkinson's disease), behavioural disorders, allergies, asthma,
 CC atherosclerosis, cardiomyopathy, angina pectoris, autoimmune diseases,
 CC retinal disease, cirrhosis, diabetes, infectious disease (bacterial,
 CC fungal, protozoal and viral e.g. human immunodeficiency virus, HIV),
 CC cancer (e.g. prostate cancer), hypertension, hypotension, multiple
 CC sclerosis, urinary retention, osteoporosis, Crohn's disease, ulcers,
 CC neurological disorders (e.g. anxiety), haemophilia or cirrhosis. They
 CC may also be used as immunogens to produce antibodies specific for the
 CC invention, and as vaccines. Further, they are useful for screening
 CC potential agonist and antagonist compounds. The sequences presented in
 CC ABG76937-ABG76956 are the human SEC1-12 and NOV1-8 proteins.

XX
 SO Sequence 381 AA;
 Query Match 99.1%; Score 2098; DB 23; Length 381;
 Best Local Similarity 99.0%; Pred. No. 5,4e-161;
 Matches 377; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MSSRIARALALVYLLHLTRIALSTCPAACHCPLAARCAAGVGVLDGCGCCVCAKOL 60
 DB 1 MSSRIARALALVYLLHLTRIALSTCPAACHCPLAARCAAGVGVLDGCGCCVCAKOL 60
 QY 61 NEDCSKTQPCDHTKGLCECNFGASSTALGICRAOSEGRPCRYNSRIYONGSEFOPNCOHQ 120
 DB 61 NEDCSKTQPCDHTKGLCECNFGASSTALGICRAOSEGRPCRYNSRIYONGSEFOPNCOHQ 120
 QY 121 CTCIDGAVGCTPLCPQELSLPNLGCPNRLVKVTGQCCCEWVCDSDSIKDMEDQDGLG 180
 DB 121 CTCIDGAVGCTPLCPQELSLPNLGCPNRLVKVTGQCCCEWVCDSDSIKDMEDQDGLG 180
 QY 181 KELGFDAEVELTRNNELIANGKRSILKRLPVFPMERIRIYNPLOGOKCTVQTSWSQCS 240
 DB 181 KELGFDAEVELTRNNELIANGKRSILKRLPVFPMERIRIYNPLOGOKCTVQTSWSQCS 240
 QY 241 KTCGTGISTRTVNDNPECRILVKEIRICEVPRPCGQPVYSLSLKKGKSKTKKSPVPRFTY 300
 DB 241 KTCGTGISTRTVNDNPECRILVKEIRICEVPRPCGQPVYSLSLKKGKSKTKKSPVPRFTY 300
 QY 301 AGCLSVKRYRKYCGSCVDGRCTPOLTRVYKMFRCEDGETFSKNVMIOSCKNVCNP 360
 DB 301 AGCLSVKRYRKYCGSCVDGRCTPOLTRVYKMFRCEDGETFSKNVMIOSCKNVCNP 360
 QY 361 HANEAPFPYRLFNDIHKFRD 381
 DB 361 HANEAPFPYRLFNDIHKFRD 381
 RESULT 10
 ABB09202
 ID ABB09202 standard; Protein: 374 AA.
 XX
 AC ABB09202;
 XX
 DT 08-JUL-2002 (first entry)
 XX
 DE HCGF CNN family protein sequence SEQ ID NO:12.
 XX
 KW Human; small CCN-like growth factor; SCGF; vulnerary; osteopathic;
 KW gene therapy; muscle wasting disease; osteoporosis; wound healing;
 KW tissue regeneration; angiogenesis.
 XX
 OS Unidentified.
 XX
 PN US2002049304-A1.
 PD 25-APR-2002.
 XX
 PF 14-MAY-2001; 2001US-0853625.
 PR 06-JUN-1995; 95US-0468847.
 PR 01-APR-1998; 98US-0053587.
 XX
 PA (HAST/) HASTINGS G A.
 PA (ADAM/) ADAMS M D.
 PI Hastings GA, Adams MD;
 DR WPI: 2002-382150/41.
 XX
 PT Novel isolated polynucleotide sequence encoding a human small CCN-like
 PT growth factor, useful for treating muscle wasting disease, and
 PT osteoporosis -
 XX
 PS Disclosure; Fig 2A-D; 33pp; English.

XX The present invention describes human small CCN-like growth factor
 CC (SCGF). SCGF has vulnary and osteopathic activities, and can be used
 CC in gene therapy. The SCGF polypeptides and polynucleotides can be used
 CC for treating muscle wasting diseases, and osteoporosis, and to stimulate
 CC wound healing and tissue regeneration, to promote angiogenesis and to
 CC stimulate proliferation of vascular smooth muscle and endothelial cell
 CC production. The present sequence represents a CCN family protein which
 CC is given in comparison with the human SCGF in the exemplification of the
 CC present invention.

XX Sequence 374 AA:

Query Match 93.2%; Score 1971.5; DB 23; Length 374;

Best Local Similarity 95.7%; Pred. No. 8.5e-151;

Matches 358; Conservative 2; Mismatches 13; Indels 1; Gaps 1;

QY 1 MSSRIARALAVYTLHLRLALSTCPACHCPLKAPGVLVBDGCGCCKVCANOL 60
 DB 1 MSSRIARALAVYTLHLRLALSTCPACHCPLKAPGVLVBDGCGCCKVCANOL 60
 QY 61 NEDCKTOPCDHTKGLKGFNFGASTALKGICRAOSBGRCEYNSRYONGESFOPNCQHO 120
 DB 61 NEDCKTOPCDHTKGLKGFNFGASTALKGICRAOSBGRCEYNSRYONGESFOPNCQHO 120
 QY 121 CTCIDGAVG-CIPLCPOELSLPNLGPVRLVKGCCCEWVCDSDIKDPMEDODGLI 179
 DB 121 CTCIDGAVG-CIPLCPOELSLPNLGPVRLVKGCCCEWVCDSDIKDPMEDODGLI 180
 QY 180 GKELGFDAEVELTRNNELIAGKRSILKRLPFGEPRILNPLDLOGKCIYOTTSMWOC 239
 DB 181 GKELGFDAEVELTRNNELIAGKRSILKRLPFGEPRILNPLDLOGKCIYOTTSMWOC 240
 QY 240 SKTCGTGISTRYTNDNPECLVETRICERPCGQVYSSILKKGKCKSTKSPVPRFT 299
 DB 241 SKTCGTGISTRYTNDNPECLVETRICERPCGQVYSSILKKGKCKSTKSPVPRFT 300
 QY 300 YAGCLSVKKYRKYKGGSCVDGRCTPOLRTYKMFRCDEGTFESKNVMIOGSKCNVNC 359
 DB 301 YAGCLSVKKYRKYKGGSCVDGRCTPOLRTYKMFRCDEGTFESKNVMIOGSKCNVNC 360
 QY 360 PHANEAAPFYRLF 373
 DB 361 PHANEAAPFYRLF 374

RESULT 11

AA90919 standard; Protein; 375 AA.

XX AAR90919;
 XX 25-MAR-2003 (updated)
 DT 25-JUN-1996 (first entry)
 XX
 DE Connective tissue growth factor-2.
 XX
 KM CTGF-2; connective tissue growth factor-2; secreted protein;
 KM cartilaginous growth; skeletal; embryo; cell growth; morphogenesis;
 XX insulin-like growth factor; fibroblast growth factor; Crg1.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..24

FT Protein /label= signal_peptide

FT Protein 25..375

XX /label= mature_protein

PN WO9601896-A1.

XX 25-JAN-1996.

XX

PF 12-JUL-1994; 94NO-US07736.
 XX
 PR 12-JUL-1994; 94NO-US07736.
 XX
 PA (HUMA-) HUMAN GENOME SCI. 18C.
 XX
 PI Adams MD, Li H;
 XX
 DR WPI: 1996-097626/10.
 DR N-PSDB; AAT12653.

PT Connective tissue growth factor-2 and DNA encoding it - useful to
 PT enhance the repair of connective and support tissue, and to enhance
 PT wound healing

PS Claim 1: Fig 1A-C; 46pp; English.

XX Connective tissue growth factor-2 (CTGF-2) is encoded by a cDNA
 CC (AAT12653) isolated from a human foetal lung cDNA library. The CTGF
 CC polypeptides are structurally and functionally related to a family
 CC of growth factors which include IGF (insulin-like growth factor),
 CC PDGF (platelet-derived growth factor), and FGF (fibroblast growth
 CC factor). CTGF-2 exhibits 89 percent identity and 93 percent similarity
 CC to Crg1. Crg1 is a growth factor-inducible immediate early gene
 CC initially identified in serum-stimulated mouse fibroblasts. It encodes
 CC a member of an emerging family of secreted proteins which are also a
 CC group of cysteine-rich proteins. This group of GFs are important for
 CC normal growth, differentiation, morphogenesis of the cartilaginous
 CC skeleton of an embryo and cell growth.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 375 AA:

Query Match 93.2%; Score 1971.5; DB 17; Length 375;

Best Local Similarity 95.7%; Pred. No. 8.5e-151;

Matches 358; Conservative 2; Mismatches 13; Indels 1; Gaps 1;

QY 1 MSSRIARALAVYTLHLRLALSTCPACHCPLKAPGVLVBDGCGCCKVCANOL 60
 DB 1 MSSRIARALAVYTLHLRLALSTCPACHCPLKAPGVLVBDGCGCCKVCANOL 60
 QY 61 NEDCKTOPCDHTKGLKGFNFGASTALKGICRAOSBGRCEYNSRYONGESFOPNCQHO 120
 DB 61 NEDCKTOPCDHTKGLKGFNFGASTALKGICRAOSBGRCEYNSRYONGESFOPNCQHO 120
 QY 121 CTCIDGAVG-CIPLCPOELSLPNLGPVRLVKGCCCEWVCDSDIKDPMEDODGLI 179
 DB 121 CTCIDGAVG-CIPLCPOELSLPNLGPVRLVKGCCCEWVCDSDIKDPMEDODGLI 180
 QY 180 GKELGFDAEVELTRNNELIAGKRSILKRLPFGEPRILNPLDLOGKCIYOTTSMWOC 239
 DB 181 GKELGFDAEVELTRNNELIAGKRSILKRLPFGEPRILNPLDLOGKCIYOTTSMWOC 240
 QY 240 SKTCGTGISTRYTNDNPECLVETRICERPCGQVYSSILKKGKCKSTKSPVPRFT 299
 DB 241 SKTCGTGISTRYTNDNPECLVETRICERPCGQVYSSILKKGKCKSTKSPVPRFT 300
 QY 300 YAGCLSVKKYRKYKGGSCVDGRCTPOLRTYKMFRCDEGTFESKNVMIOGSKCNVNC 359
 DB 301 YAGCLSVKKYRKYKGGSCVDGRCTPOLRTYKMFRCDEGTFESKNVMIOGSKCNVNC 360
 QY 360 PHANEAAPFYRLF 373
 DB 361 PHANEAAPFYRLF 374

RESULT 12

AA931620 standard; Protein; 375 AA.

XX AAY31620;

XX 02-NOV-1999 (first entry)

XX Human CTGF-2.
 DE Connective tissue growth factor-2; CTGF-2; wound healing; bone disorder;
 XX skin disorder; acne; burn; UV damage; stabilisation; tissue implant.
 KW Homo sapiens.
 XX
 OS
 FH Key Location/Qualifiers
 FT Peptide 1..24
 FT /label= Signal_peptide
 FT Protein 25..375
 FT Misc-difference 268
 FT /note= "Cys encoded by ICT"
 XX
 PN US945300-A.
 XX
 PD 31-AUG-1999.
 XX
 PF 02-JUN-1995; 95US-0459101.
 XX
 PR 02-JUN-1995; 95US-0459101.
 PR 12-JUL-1994; 94WO-US07736.
 XX
 PA (ADAM/) ADAMS M D.
 PA (LITH/) LI H.
 XX
 PI Adams MD, LI H;
 XX
 DR WPI: 1999-508171/42.
 DR N-PSDB: AA211720.
 XX
 PT Polynucleotides encoding growth factor polypeptides useful for
 PT enhancing the repair of connective tissue and support tissue
 XX
 PS Claim 1; Fig 1; 20pp; English.
 XX
 CC This sequence represents human connective tissue growth factor-2
 CC (CTGF-2). CTGF-2 cDNA was isolated from a cDNA library derived from
 CC human foetal lung. In one instance, the cDNA was cloned into a
 CC baculovirus expression vector, having first been amplified and modified
 CC via PCR using primers AA211721 and AA211722. In another instance, the
 CC cDNA was cloned into a COS cell expression vector, with prior
 CC amplification and modification using PCR primers AA211723 and AA211724.
 CC CTGF-2 is structurally and functionally related to a family of growth
 CC factors which include IGF (insulin-like growth factor), PDGF
 CC (platelet-derived growth factor) and FGF (fibroblast growth factor). This
 CC emerging family of cysteine-rich secreted proteins are important for
 CC normal growth, differentiation, morphogenesis of the cartilaginous
 CC skeleton of an embryo and cell growth. Their functions also include wound
 CC healing, tissue repair, implant fixation and stimulating increased bone
 CC mass. CTGF-2 may be used to enhance the repair of connective tissue and
 CC support tissue and can therefore treat skin disorders e.g., acne, aging,
 CC UV damage or burns. CTGF-2 can be used to promote the attachment,
 CC fixation and stabilisation of tissue implants inserted during
 CC reconstructive surgery, and can be used to enhance the healing of
 CC external wounds. It can be used in the treatment of injured or depleted
 CC bone as it promotes the growth of connective tissue, bone and cementum
 CC and stimulates protein and collagen synthesis.
 CC
 SQ Sequence 375 AA;
 XX
 Query Match 93.2%; Score 1971.5; DB 20; Length 375;
 Best Local Similarity 95.7%; Pred. No. 8.5e-151;
 Matches 358; Conservative 2; Mismatches 13; Indels 1; Gaps 1;
 QY 1 MSSRIARALAVVTLHLTRIALSTCPACHCPLEAPKCAPGVGLVDGGCCCKVCAKOL 60
 DB 1 MSSRIARALAVVTLHLTRIALSTCPACHCPLEAPKCAPGVGLVDGGCCCKVCAKOL 60
 QY 61 NEQSKRQPDHDKRGLECNAGSSTALKGICRAOSEPPEYNSRIRYONESFOPNCOHQ 120
 DB 61 NEQSKRQPDHDKRGLECNAGSSTALKGICRAOSEPPEYNSRIRYONESFOPNCKHQ 120

QY 121 CTCIDGAVG-CIPLCPQELSLPNTGCPNPRVLKVTGOCCEMWCDSDSIRKDPMDQDGL 179
 DB 121 CTCIGMRGACIPLCPQELSLPNTGCPNPRVLKVTGOCCEMWCDSDSIRKDPMDQDGL 180
 QY 180 GKELGFPDASEVELTRNNELIYVAGKRSILKRLPVEGMERIRIYNPLOGKCIYQTTSMSC 239
 DB 181 GKGLGFPDASEVELTRNNELIYVAGKRSILKRLPVEGMERIRIYNPLOGKCIYQTTSMSC 240
 QY 240 SKTCGTGTSTRTVTDNPECRILVETRICENAPCCQPIYSSLKAKKSKTRKSPVAF 299
 DB 241 SKTCGTGTSTRTVTDNPECRILVETRICENAPCCQPIYSSLKAKKSKTRKSPVAF 300
 QY 300 YAGCLSVKKYRKYRCGSCVDGRCCTPOLTRIVKRRFRCEDEFTSKNNVMOQKCNVC 359
 DB 301 YAGCLSVKKYRKYRCGSCVDGRCCTPOLTRIVKRRFRCEDEFTSKNNVMOQKCNVC 360
 QY 360 PHANEAAFPFYRLF 373
 DB 361 PHANEAAFPFYRLF 374
 RESULT 13
 ID AAE18108 standard; Protein: 375 AA.
 AC AAE18108;
 DT 07-MAY-2002 (first entry)
 DE Human alternative connective tissue growth factor-2 (CTGF-2).
 XX
 KW Human: angiogenesis; connective tissue growth factor-2; CTGF-2; tumour;
 KW ischaemia; restenosis; tissue repair; wound healing; congenital defect;
 KW cardiovascular disease; atherosclerosis; heart failure; trauma;
 KW burns; osteoporosis; periodontal disease; liver failure; transquillizer;
 KW vulnery; cosmetic plastic surgery; vasotropic; hepatotropic; ulcer;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO200204480-A2.
 PD 17-JAN-2002.
 XX
 PF 11-JUL-2001; 2001WO-US21799.
 XX
 PR 11-JUL-2000; 2000US-217402P.
 PR 18-MAY-2001; 2001US-291642P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (TRGE) TRANSGENE SA.
 XX
 PI LI H, Adams MD, Calenda V, Fataccioli V;
 XX
 DR WPI: 2002-171698/22.
 DR N-PSDB: AAD29099.
 PT Stimulating angiogenesis in a mammal preferably human having ischemia
 PT or restenosis or is treated for limb revascularization, by
 PT administering connective tissue growth factor-2 polypeptide or
 PT polynucleotide
 XX
 PS Disclosure; Fig 11; 13pp; English.
 XX
 CC The present invention relates to a method for stimulating angiogenesis in
 CC a mammal. The method comprises administering a polynucleotide encoding
 CC connective tissue growth factor-2 (CTGF-2) or an active fragment or its
 CC derivative. The method is useful for stimulating angiogenesis in a mammal
 CC preferably human having ischemia or restenosis or is treated for limb
 CC revascularisation which is leg or arm. The invention is useful for
 CC inhibiting tumour growth, where angiogenesis is utilised for enhancing
 CC the repair of connective and support tissue, promoting the attachment,

CC fixation and stabilisation of tissue implants and enhancing wound
 CC healing, hence is useful for treating cardiovascular disease e.g.
 CC atherosclerosis, reperfusion injury such as heart failure, angina,
 CC ischaemia; and is also used to differentiate, proliferate and attract
 CC cells leading to regeneration of tissues which is utilised to repair
 CC replace or protect tissue damaged by congenital defects, trauma (burns,
 CC ulcer, etc), age, disease (e.g. osteoporosis, periodontal disease,
 CC liver failure), surgery including cosmetic plastic surgery. The present
 CC sequence is human alternative CTGF-2. CTGF-2 gene is useful in gene
 CC therapy.

XX Sequence 375 AA:

Query Match 93.2%; Score 1971.5; DB 23; Length 375;

Best Local Similarity 95.7%; Pred. No. 8.5e-151;
 Matches 358; Conservative 2; Mismatches 13; Indels 1; Gaps 1;

QY 1 MSSRIARALAVYTLHLTRALSTCPACHCPLKAPKAPGVGLVBDGCGCKVCAKOL 60

DB 1 MSSRIARALAVYTLHLTRALSTCPACHCPLKAPKAPGVGLVBDGCGCKVCAKOL 60

QY 61 NEDCSKTQPCDHTKGLKCNFAGASTALKGICRAOSGRCPEYNSRIYQNGESFOPNCKHQ 120

DB 61 NEDCSKTQPCDHTKGLKCNFAGASTALKGICRAOSGRCPEYNSRIYQNGESFOPNCKHQ 120

QY 121 CTCIDGAVG-CIPCLPQELSLPNIGCPNRLVVTGGCCBEMWCDEDSIKDPMEDDGLL 179

DB 121 CTCIDGAVG-CIPCLPQELSLPNIGCPNRLVVTGGCCBEMWCDEDSIKDPMEDDGLL 180

QY 180 GKEIGDASVEYELTRNNELIYAVGKSLKRLPVFGMEPRLLYNPLOGKCIYVOTTSWQC 239

DB 181 GKEIGDASVEYELTRNNELIYAVGKSLKRLPVFGMEPRLLYNPLOGKCIYVOTTSWQC 240

QY 240 SKTCGTGISTRTVNDNPECKRLVETRIKCEVRPCGOPYSSILKKKKSKTKKSPPEVRF 299

DB 241 SKTCGTGISTRTVNDNPECKRLVETRIKCEVRPCGOPYSSILKKKKSKTKKSPPEVRF 300

QY 300 YAGCLSVKTKRPRYCGSCVDGRCTPOLTRTYKMRFCDEDETFSKNVMYIOSCKNYNC 359

DB 301 YAGCLSVKTKRPRYCGSCVDGRCTPOLTRTYKMRFCDEDETFSKNVMYIOSCKNYNC 360

QY 360 PHANEAFPPYRLF 373

DB 361 PHANEAFPPYRLF 374

RESULT 14

AAR25565 standard; Protein; 379 AA.

XX AAR25565;

XX 25-MAR-2003 (updated)

DT 18-JAN-1993 (first entry)

XX Beta-IG-M1.

XX Transforming growth factor beta; induced; CEF-10; v-src; chicken;

KM embryo; fibroblasts; TGF-beta.

XX Mus musculus.

PN EP495674-A2.

PD 22-JUL-1992.

PF 17-JAN-1992; 92EP-0300429.

PR 18-JAN-1991; 91US-0642991.

XX 10-JAN-1992; 92US-0816270.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

PI Brunner AM, Chinn J, Neubauer MG, Purchio AF;
 XX WPI; 1992-243508/30.
 DR N-PSDB; AAQ26421.
 XX TGF-beta induced gene family - encodes proteins involved in
 PT growth and differentiation effects of TGF-beta-1
 PS Claim 2; Fig 1; 35pp; English.

CC The protein sequence was deduced from the DNA sequence obtd. by
 CC screening a cDNA library made from Akr-2B mouse cells induced with
 CC TGF-beta1 and cyclohexamide with two probes from untreated Akr-2B
 CC mRNA and Akr-2B mRNA from cells treated with cyclohexamide and TGF-
 CC beta1. The proteins encoded by hybridising colonies (beta-IG-M1 and
 CC beta-IG-M2) contain 38 Cys residues and are induced by TGF-beta1.
 CC Beta-IG-M1 displays 80 percent homology to the CEF-10 protein
 CC induced by v-src in chicken embryo fibroblasts and is identical
 CC to the protein encoded by cyf61, an immediate early response gene
 CC induced in quiescent Balb 3T3 cells by serum treatment. Residues
 CC 49-56 of beta-IG-M1 conform to the GCGCCXXC motif reported in the
 CC amino half of insulin-like growth factor (IGF) binding proteins.
 CC The C-terminal Cys rich region of beta-IG-M1, -M2 and CEF-10 contain
 CC an amino acid sequence with strong homology to a motif found near the
 CC C-terminal of the malarial circumsporozoite (CS) protein, which is
 CC highly conserved among all species of malarial parasites sequenced
 CC to date (designated region II). This motif is also found in
 CC other proteins which have cell adhesive properties that mediate
 CC cell-cell and cell-extracellular matrix interactions, such as
 CC perlecan, thrombospondin, and TRAP. The proteins encoded by
 CC TGF-beta induced genes are likely to be involved in mediation of
 CC the biological effects of TGF-beta relating to cell growth and
 CC differentiation. See also AAR25566.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 379 AA:

Query Match 91.2%; Score 1929; DB 13; Length 379;

Best Local Similarity 90.9%; Pred. No. 2.3e-147;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSSRIARALAVYTLHLTRALSTCPACHCPLKAPKAPGVGLVBDGCGCKVCAKOL 60

DB 1 MSSRIARALAVYTLHLTRALSTCPACHCPLKAPKAPGVGLVBDGCGCKVCAKOL 60

QY 61 NEDCSKTQPCDHTKGLKCNFAGASTALKGICRAOSGRCPEYNSRIYQNGESFOPNCKHQ 120

DB 61 NEDCSKTQPCDHTKGLKCNFAGASTALKGICRAOSGRCPEYNSRIYQNGESFOPNCKHQ 120

QY 121 CTCIDGAVG-CIPCLPQELSLPNIGCPNRLVVTGGCCBEMWCDEDSIKDPMEDDGLL 180

DB 121 CTCIDGAVG-CIPCLPQELSLPNIGCPNRLVVTGGCCBEMWCDEDSIKDPMEDDGLL 180

QY 181 KEGIGDASVEYELTRNNELIYAVGKSLKRLPVFGMEPRLLYNPLOGKCIYVOTTSWQC 238

DB 179 --LGLDASEVELTRNNELIYAVGKSLKRLPVFGMEPRLLYNPLOGKCIYVOTTSWQC 236

QY 239 CSKTCGTGISTRTVNDNPECKRLVETRIKCEVRPCGOPYSSILKKKKSKTKKSPPEVRF 298

DB 237 CSKTCGTGISTRTVNDNPECKRLVETRIKCEVRPCGOPYSSILKKKKSKTKKSPPEVRF 296

QY 299 TYAGCLSVKTKRPRYCGSCVDGRCTPOLTRTYKMRFCDEDETFSKNVMYIOSCKNYNC 358

DB 297 TYAGCLSVKTKRPRYCGSCVDGRCTPOLTRTYKMRFCDEDETFSKNVMYIOSCKNYNC 356

QY 359 CPANEAFPPYRLFNDIHKFRD 381

DB 357 CPANEAFPPYRLFNDIHKFRD 379

RESULT 15

AAE05920 standard; Protein; 379 AA.

XX AA05920;
 AC 24-SEP-2001 (first entry)
 DT Mouse cysteine-rich protein (Cyr61).
 DE Cysteine-rich protein; Cyr61; extracellular matrix signalling molecule;
 KM fibroblast secreted protein; fisp12; connective tissue growth factor;
 KM CTGF; ECM: cell adhesion; cell migration; fibroblast cell proliferation;
 KM angiogenesis; wound healing; integrin receptor; atherosclerosis; tumour;
 KM heart disease; fibrosis; gene therapy; mouse.
 XX Mus musculus.
 OS
 XX
 FH Key Location/Qualifiers
 FT Region 164..226
 FT /note="Cysteine free region"
 FT 224..240
 FT Domain /note="Domain III"
 XX
 XX WO200155210-A2.
 PN
 XX
 PD 02-AUG-2001.
 XX
 PF 31-JAN-2001; 2001WO-US03267.
 XX
 PR 31-JAN-2000; 2000US-0495448.
 PR 15-MAY-2000; 2000US-0204364.
 PR 06-OCT-2000; 2000US-0238705.
 XX
 PA (MUNI-) MUNIN CORP.
 XX
 PI Lau LF, Yeung C, Greenspan JA;
 XX
 DR WPI: 2001-465561/50.
 DR N-PSDB: AAD11220.
 XX
 PT Novel human cysteine-rich protein 61 (Cyr61) fragment useful in methods
 PT for screening for modulators of cell adhesion, fibroblast cell
 PT proliferation, angiogenesis and cell migration
 XX
 PS Example 1; Fig 1; 186pp; English.
 XX
 CC The invention relates to extracellular matrix (ECM) signalling
 CC molecules involved in cellular response to growth factors. More
 CC particularly the invention is directed to cysteine-rich protein
 CC (Cyr61), and Cyr61-related proteins such as fibroblast secreted
 CC protein (fisp12) and connective tissue growth factor (CTGF) and
 CC nucleic acid molecules encoding such proteins. The polypeptides
 CC of the invention are members of cysteine-rich secreted protein
 CC family. Human Cyr61 fragment is useful in methods for screening
 CC modulators of cell adhesion, cell migration, fibroblast cell
 CC proliferation, angiogenesis, wound healing and Cyr61-integrin
 CC receptor interaction. Modulator of Cyr61-integrin alphavbeta3
 CC interaction is used for the preparation of a medicament for the
 CC treatment of atherosclerosis, heart disease, tumour metastasis,
 CC fibrosis, tumour growth, disorders associated with inadequate
 CC angiogenesis; aberrant granulation tissue development; aberrant
 CC fibroblast growth and wounds. Polynucleotides of the invention
 CC are useful in gene therapy. The present sequence is mouse Cyr61
 CC protein.
 XX
 SO Sequence 379 AA;

Query Match 91.2%; Score 1929; DB 22; Length 379;
 Best local similarity 90.9%; Pred. No. 2.3e-147;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSSRIARALAVTLHLTRALSTCPRACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60
 DB 1 MSSSTFRLTAAVTLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60

QY 61 NEDCSKTQPCDHTKGLGECNFGASSTALGICRAQSEGRCEYNSRIYONGESFOPNCOHQ 120
 DB 61 NEDCSKTQPCDHTKGLGECNFGASSTALGICRAQSEGRCEYNSRIYONGESFOPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCQOELSLPMLGCPNRLVAVYTGOCCEWYCDDESDIKDPHEDDGLG 180
 DB 121 CTCIDGAVGCIPLCQOELSLPMLGCPNRLVAVYTGOCCEWYCDDESDIKDSLDDDDL-- 178
 QY 181 KELGPDASEVELTRNNELIANGKRSIKRLPYEGMEPRILYNPL--QGOKCIYOTSMQ 238
 DB 179 --LGIDASEVELTRNNELIANGKSSLRPLPYGTEPRVLEPLHAHQKCIYOTSMQ 236
 QY 239 CSKTGCTGISTRTVNDNEBECRLVKEPRICEVAPCGQPYSSLSKKGKCSKTKKSPDYRF 298
 DB 237 CSKSGCTGISTRTVNDNEBECRLVKEPRICEVAPCGQPYSSLSKKGKCSKTKKSPDYRF 296
 QY 299 TYAGLSVKKYRPKYGSCVDGRCTPOLTRTKMRFRCEDEGETFSKNVMIQSKCNYN 358
 DB 297 TYAGSSVKKYRPKYGSCVDGRCTPOLTRTKMRFRCEDEGETFSKNVMIQSKCNYN 356
 QY 359 CPHANEAPFPYRLFNDIHKPRD 381
 DB 357 CPHNEASFRLYSLFNDIHKPRD 379

Search completed: August 5, 2003, 14:05:12
 Job time : 40.1026 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:03:52 ; Search time 14.0368 Seconds
(without alignments)
1148.438 Million cell updates/sec

Title: US-09-495-448a-4

Perfect score: 2116

Sequence: 1 MSSRIARALALVTLHLTR.....ANEAAPFYRLFNDRFRD 381

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTOUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2116	100.0	381	4	US-09-142-569-4
2	2106	99.5	381	4	US-09-348-815-2
3	1971.5	93.2	374	1	US-08-468-847B-12
4	1964.5	92.8	375	2	US-08-459-101A-2
5	1929	91.2	379	1	US-08-468-847B-11
6	1929	91.2	379	4	US-09-142-569-2
7	1695	80.1	375	1	US-08-468-847B-13
8	953	45.0	347	4	US-09-582-337-2
9	952.5	45.0	348	4	US-09-292-036-3
10	948.5	44.8	348	1	US-08-468-847B-15
11	948.5	44.8	347	4	US-09-142-569-6
12	948	44.8	347	4	US-09-187-478-2
13	948	44.8	347	4	US-09-292-036-2
14	947	44.8	349	1	US-08-167-628-2
15	947	44.8	349	1	US-08-386-680-2
16	947	44.8	349	1	US-08-459-717-2
17	947	44.8	349	1	US-08-712-302-2
18	947	44.8	349	2	US-08-880-031-2
19	947	44.8	349	3	US-09-054-358-2
20	947	44.8	349	3	US-09-097-119-2
21	947	44.8	349	3	US-09-054-274-2
22	947	44.8	349	3	US-09-080-715-2
23	947	44.8	349	3	US-09-056-704-2
24	947	44.8	349	4	US-09-292-036-4
25	947	44.8	349	4	US-09-253-316-26
26	947	44.8	349	4	US-09-142-569-8
27	947	44.8	349	4	US-09-461-688-2

28	947	44.8	349	5	PCR-US96-08140-2	Sequence 2, Appl1
29	938.5	44.4	348	1	US-08-468-847B-14	Sequence 14, Appl
30	851.5	40.2	351	1	US-08-468-847B-16	Sequence 16, Appl
31	827.5	39.1	357	1	US-08-468-847B-17	Sequence 17, Appl
32	827.5	39.1	357	1	US-09-253-316-25	Sequence 25, Appl
33	767	36.2	367	4	US-09-182-145-4	Sequence 4, Appl1
34	767	36.2	367	4	US-09-182-145-8	Sequence 8, Appl1
35	766	36.2	367	4	US-09-182-145-7	Sequence 7, Appl1
36	766	36.2	367	4	US-09-182-145-22	Sequence 22, Appl
37	760.5	35.9	345	4	US-09-182-145-3	Sequence 3, Appl1
38	760.5	35.9	345	4	US-09-182-145-6	Sequence 6, Appl1
39	759.5	35.9	345	4	US-09-182-145-5	Sequence 5, Appl1
40	759.5	35.9	345	4	US-09-182-145-21	Sequence 21, Appl1
41	758	35.8	367	4	US-09-182-145-12	Sequence 12, Appl
42	754.5	35.7	345	4	US-09-182-145-11	Sequence 11, Appl
43	613	28.0	339	4	US-09-182-145-36	Sequence 36, Appl
44	613	28.0	354	4	US-09-182-145-37	Sequence 37, Appl
45	613	29.0	354	4	US-09-253-316-2	Sequence 2, Appl1

ALIGNMENTS

```
RESULT 1
US-09-142-569-4
; Sequence 4, Application US/09142569
; Patent No. 6413735
GENERAL INFORMATION:
APPLICANT: Lau, Lester F.
TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Noole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/142,569
FILING DATE: 02-Apr-1999
CLASSIFICATION: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28758/33766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6500
TELEFAX: 312/474-0448
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 381 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: "Human Cyt61 amino acid sequence"
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-142-569-4
Query Match 100.0%; Score 2116; DB 4; Length 381;
Best Local Similarity 100.0%; Pred. No. 7.8e-177;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MSSRIARALALVTLHLTRALSTCPAACHCPLEAPKARGVGLVRDGCCKRYCAKOL 60
|||||
```

Db 1 MSSRIARALALVYTLHLTRIALSTCPACCHCLEAPKCAPGVGLVDDGCCCKVCAKOL 60
QY 61 NEDCSKTOPCDHDKKGLGCNFGASTALKGICRAOSGSRPCEVNSRIYONGESPQPCNOHQ 120
Db 61 NEDCSKTOPCDHDKKGLGCNFGASTALKGICRAOSGSRPCEVNSRIYONGESPQPCNOHQ 120
QY 121 CTCIDGAVGCIPLCPDELSPNLGCPNRLVKTGOCCEWVDEDSIKPMEDODGLLG 180
Db 121 CTCIDGAVGCIPLCPDELSPNLGCPNRLVKTGOCCEWVDEDSIKPMEDODGLLG 180
QY 181 KELGPDASEVELTRNNELLAVGKRSILKRLPVFGMEPRILYNPLGQCKIVQTTSSQCS 240
Db 181 KELGPDASEVELTRNNELLAVGKRSILKRLPVFGMEPRILYNPLGQCKIVQTTSSQCS 240
QY 241 KTCGTGISTRVTDNPECRVLTKEITRICEVRPGOPVYSSLKKGKCKSKTKKSPVRFY 300
Db 241 KTCGTGISTRVTDNPECRVLTKEITRICEVRPGOPVYSSLKKGKCKSKTKKSPVRFY 300
QY 301 AGCLSVKXKYPKYGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMIQSCKNYCP 360
Db 301 AGCLSVKXKYPKYGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMIQSCKNYCP 360
QY 361 HANEAPFPYRLFNDIHKFRD 381
Db 361 HANEAPFPYRLFNDIHKFRD 381

RESULT 2

US-09-348-815-2
Sequence 2, Application US/09348815
Patent No. 6534630
GENERAL INFORMATION:
APPLICANT: LI, HAODONG
ADAMS, MARK D
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR-2
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: HUMAN GENOME-SCIENCES, INC.
STREET: 9410 KEY WEST AVENUE
CITY: ROCKVILLE
STATE: MD
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/348,815
FILING DATE: 08-Jul-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: JONATHAN L. KLEIN
REGISTRATION NUMBER: 41,119
REFERENCE/DOCKET NUMBER: PFI26PID1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 381 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-348-815-2

Query Match 99.5%; Score 2106; DB 4; Length 381;
Best local Similarity 99.5%; Pred. No. 5,8e-176;
Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MSSRIARALALVYTLHLTRIALSTCPACCHCLEAPKCAPGVGLVDDGCCCKVCAKOL 60

Db 1 MSSRIARALALVYTLHLTRIALSTCPACCHCLEAPKCAPGVGLVDDGCCCKVCAKOL 60
QY 61 NEDCSKTOPCDHDKKGLGCNFGASTALKGICRAOSGSRPCEVNSRIYONGESPQPCNOHQ 120
Db 61 NEDCSKTOPCDHDKKGLGCNFGASTALKGICRAOSGSRPCEVNSRIYONGESPQPCNOHQ 120
QY 121 CTCIDGAVGCIPLCPDELSPNLGCPNRLVKTGOCCEWVDEDSIKPMEDODGLLG 180
Db 121 CTCIDGAVGCIPLCPDELSPNLGCPNRLVKTGOCCEWVDEDSIKPMEDODGLLG 180
QY 181 KELGPDASEVELTRNNELLAVGKRSILKRLPVFGMEPRILYNPLGQCKIVQTTSSQCS 240
Db 181 KELGPDASEVELTRNNELLAVGKRSILKRLPVFGMEPRILYNPLGQCKIVQTTSSQCS 240
QY 241 KTCGTGISTRVTDNPECRVLTKEITRICEVRPGOPVYSSLKKGKCKSKTKKSPVRFY 300
Db 241 KTCGTGISTRVTDNPECRVLTKEITRICEVRPGOPVYSSLKKGKCKSKTKKSPVRFY 300
QY 301 AGCLSVKXKYPKYGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMIQSCKNYCP 360
Db 301 AGCLSVKXKYPKYGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMIQSCKNYCP 360
QY 361 HANEAPFPYRLFNDIHKFRD 381
Db 361 HANEAPFPYRLFNDIHKFRD 381

RESULT 3

US-08-468-847B-12
Sequence 12, Application US/08468847B
Patent No. 5780263
GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,847B
FILING DATE: 6 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1744
TELEFAX: 201-994-1700
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 374 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-468-847B-12

Query Match 93.2%; Score 1971.5; DB 1; Length 374;

Best Local Similarity 95.7%; Pred. No. 3.1e-164;
Matches 358; Conservative 2; Mismatches 13; Indels 1; Gaps 1;

```
QY 1 MSSRRARLAVITLHLITRLALSTCPACHPLEAPCAPGVGLVRGCGCCCKVCAROL 60
  |||||
Db 1 MSSRRARLAVITLHLITRLALSTCPADCHCPLAPCAPGVGLVRGCGCCCKVCAROL 60
  |||||
QY 61 NEDSKTOPCDHTKGLKLECNFGASSALGICRAOSEGRPCENSRITYNGESFOPNCQHO 120
  |||||
Db 61 NEDCKTQPCDHTKGLKLECNFGASSALGICRAOSEGRPCENSRITYNGESFOPNCQHO 120
  |||||
QY 121 CTCIDGAVG-CIPICPOELSLPNLGCNPNRLVKTGQCCCEWVCDEDSIKDPMEDQDGL 179
  |||||
Db 121 CTCIGWRGACIPICPOELSLPNLGCNPNRLVKTGQCCCEWVCDEDSIKDPMEDQDGL 180
  |||||
QY 180 GKELGFDASEVELFRNNELIANGKRSILKRLPVFGMEPRILYNPLQGGKCIYQTSWSQC 239
  |||||
Db 181 GKELGFDASEVELFRNNELIANGKRSILKRLPVFGMEPRILYNPLQGGKCIYQTSWSQC 240
  |||||
QY 240 SKTGTGISTVTNDNPECRILVKEIRICEVRPCGQPVYSSILKGGKSKTKKSPPEVRF 299
  |||||
Db 241 SKTGTGISTVTNDNPECRILVKEIRICEVRPCGQPVYSSILKGGKSKTKKSPPEVRF 300
  |||||
QY 300 YAGCLSVKRYPRKTCGSCVDGRCTPOLTRTVKMFRCDEDETSKNVMYIOSKCNVNC 359
  |||||
Db 301 YAGCLSVKRYPRKTCGSCVDGRCTPOLTRTVKMFRCDEDETSKNVMYIOSKCNVNC 360
  |||||
QY 360 PHANEAPFERYRLF 373
  |||||
Db 361 PHANEAPFERYRLF 374
  |||||
```

RESULT 4

US-08-459-101A-2
Sequence 2, Application US/08459101A

Patent No. 5945300

GENERAL INFORMATION:

APPLICANT: LI, ET AL.

TITLE OF INVENTION: Connective Tissue Growth Factor-2

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESSES:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

ADDRESSEE: CECCHI, STEWART & OLSTEIN

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/459,101A

FILING DATE: June 2, 1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/07736

FILING DATE: 12 JUL 94

ATTORNEY/AGENT INFORMATION:

NAME: FERRARO, GREGORY D.

REGISTRATION NUMBER: 36,134

REFERENCE/DOCKET NUMBER: 325800-317

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 375 AMINO ACIDS

TYPE: AMINO ACID

STRANDEDNESS:

TOPOLOGY: LINEAR

MOLECULE TYPE: PROTEIN
US-08-459-101A-2

Query Match 92.8%; Score 1964.5; DB 2; Length 375;
Best Local Similarity 95.5%; Pred. No. 1.3e-163;
Matches 357; Conservative 2; Mismatches 14; Indels 1; Gaps 1;

```
QY 1 MSSRRARLAVITLHLITRLALSTCPACHPLEAPCAPGVGLVRGCGCCCKVCAROL 60
  |||||
Db 1 MSSRRARLAVITLHLITRLALSTCPADCHCPLAPCAPGVGLVRGCGCCCKVCAROL 60
  |||||
QY 61 NEDSKTOPCDHTKGLKLECNFGASSALGICRAOSEGRPCENSRITYNGESFOPNCQHO 120
  |||||
Db 61 NEDCKTQPCDHTKGLKLECNFGASSALGICRAOSEGRPCENSRITYNGESFOPNCQHO 120
  |||||
QY 121 CTCIDGAVG-CIPICPOELSLPNLGCNPNRLVKTGQCCCEWVCDEDSIKDPMEDQDGL 179
  |||||
Db 121 CTCIGWRGACIPICPOELSLPNLGCNPNRLVKTGQCCCEWVCDEDSIKDPMEDQDGL 180
  |||||
QY 180 GKELGFDASEVELFRNNELIANGKRSILKRLPVFGMEPRILYNPLQGGKCIYQTSWSQC 239
  |||||
Db 181 GKELGFDASEVELFRNNELIANGKRSILKRLPVFGMEPRILYNPLQGGKCIYQTSWSQC 240
  |||||
QY 240 SKTGTGISTVTNDNPECRILVKEIRICEVRPCGQPVYSSILKGGKSKTKKSPPEVRF 299
  |||||
Db 241 SKTGTGISTVTNDNPECRILVKEIRICEVRPCGQPVYSSILKGGKSKTKKSPPEVRF 300
  |||||
QY 300 YAGCLSVKRYPRKTCGSCVDGRCTPOLTRTVKMFRCDEDETSKNVMYIOSKCNVNC 359
  |||||
Db 301 YAGCLSVKRYPRKTCGSCVDGRCTPOLTRTVKMFRCDEDETSKNVMYIOSKCNVNC 360
  |||||
QY 360 PHANEAPFERYRLF 373
  |||||
Db 361 PHANEAPFERYRLF 374
  |||||
```

RESULT 5

US-08-468-847B-11
Sequence 11, Application US/08468847B

Patent No. 5780263

GENERAL INFORMATION:

APPLICANT: Haslings, Gregg A. and Adams, Mark D.

TITLE OF INVENTION: Human CCN-Like Growth Factor

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESSES:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

ADDRESSEE: CECCHI, STEWART & OLSTEIN

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/468,847B

FILING DATE: 6 June 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.

REGISTRATION NUMBER: 33,073

REFERENCE/DOCKET NUMBER: 325800-442

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 379 AMINO ACIDS
 TYPE: AMINO ACID
 STRANDEDNESS:
 TOPOLOGY: LINEAR
 MOLECULE TYPE: PROTEIN
 US-08-468-847B-11

Query Match 91.2%; Score 1929; DB 1; Length 379;
 Best Local Similarity 90.9%; Pred. No. 1.6e-160;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSSRIARALAVLTLLHLRLALSTCPACCHCPLEAPKAPGVGLVYRDGCGCKVCARQL 60
 DB 1 MSSSTRITLAVATLTLHLRLALSTCPACCHCPLEAPKAPGVGLVYRDGCGCKVCARQL 60
 QY 61 NEDCSKTOPCDHRTKGLCECNFGASSTALKGICRAOSGRCPEYNSRIYONGESFOPNCQHQ 120
 DB 61 NEDCSKTOPCDHRTKGLCECNFGASSTALKGICRAOSGRCPEYNSRIYONGESFOPNCQHQ 120
 QY 121 CTCIDAVGCIPLCPDELSPNLGCPNPLVYKYGOCCEWVDEDSIKDPMDODGLG 180
 DB 121 CTCIDAVGCIPLCPDELSPNLGCPNPLVYKYGOCCEWVDEDSIKDPMDODGLG 180
 QY 181 KELGFASVEELTRNNELTAVGKRSKRLPVFGMPRLIYNPL--QGOKCIYQTSMSQ 238
 DB 179 --LGDASEVELTRNNELTAVGKSSSLKRLPVFGTEPRVLFNPLHAHGOKCIYQTSMSQ 236
 QY 239 CSKTCGTGISTRYTNDNPECLVETRICVPRPCGQPVYSSLKKGKCKSTKKSPEPVRF 298
 DB 237 CSKSCGTGISTRYTNDNPECLVETRICVPRPCGQPVYSSLKKGKCKSTKKSPEPVRF 296
 QY 299 TYAGCISVKKYRKRYGSCVDGRCCTPLQTRIVKMFRCDEGEMFSKNVMIOSCKCNYN 358
 DB 297 TYAGCISVKKYRKRYGSCVDGRCCTPLQTRIVKMFRCDEGEMFSKNVMIOSCKCNYN 356
 QY 359 CPHANEAPFERYRLFNDIHKFRD 381
 DB 357 CPHNEASFRLYSLFNDIHKFRD 379

RESULT 6
 US-09-142-569-2
 ; Sequence 2, Application US/09142569
 ; Patent No. 6413735
 ; GENERAL INFORMATION:
 ; APPLICANT: Lau, Lester F.
 ; TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
 ; NUMBER OF SEQUENCES: 17
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/142.569
 ; FILING DATE: 02-Apr-1999
 ; CLASSIFICATION: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Clough, David W.
 ; REGISTRATION NUMBER: 36,107
 ; REFERENCE/DOCKET NUMBER: 28738/33766
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312/474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
 LENGTH: 379 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: misc-feature
 OTHER INFORMATION: "Mouse Cyr61 amino acid sequence"
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-142-569-2

QY 1 MSSRIARALAVLTLLHLRLALSTCPACCHCPLEAPKAPGVGLVYRDGCGCKVCARQL 60
 DB 1 MSSSTRITLAVATLTLHLRLALSTCPACCHCPLEAPKAPGVGLVYRDGCGCKVCARQL 60
 QY 61 NEDCSKTOPCDHRTKGLCECNFGASSTALKGICRAOSGRCPEYNSRIYONGESFOPNCQHQ 120
 DB 61 NEDCSKTOPCDHRTKGLCECNFGASSTALKGICRAOSGRCPEYNSRIYONGESFOPNCQHQ 120
 QY 121 CTCIDAVGCIPLCPDELSPNLGCPNPLVYKYGOCCEWVDEDSIKDPMDODGLG 180
 DB 121 CTCIDAVGCIPLCPDELSPNLGCPNPLVYKYGOCCEWVDEDSIKDPMDODGLG 180
 QY 181 KELGFASVEELTRNNELTAVGKRSKRLPVFGMPRLIYNPL--QGOKCIYQTSMSQ 238
 DB 179 --LGDASEVELTRNNELTAVGSSSLKRLPVFGTEPRVLFNPLHAHGOKCIYQTSMSQ 236
 QY 239 CSKTCGTGISTRYTNDNPECLVETRICVPRPCGQPVYSSLKKGKCKSTKKSPEPVRF 298
 DB 237 CSKSCGTGISTRYTNDNPECLVETRICVPRPCGQPVYSSLKKGKCKSTKKSPEPVRF 296
 QY 299 TYAGCISVKKYRKRYGSCVDGRCCTPLQTRIVKMFRCDEGEMFSKNVMIOSCKCNYN 358
 DB 297 TYAGCISVKKYRKRYGSCVDGRCCTPLQTRIVKMFRCDEGEMFSKNVMIOSCKCNYN 356
 QY 359 CPHANEAPFERYRLFNDIHKFRD 381
 DB 357 CPHNEASFRLYSLFNDIHKFRD 379

RESULT 7
 US-08-468-847B-13
 ; Sequence 13, Application US/08468847B
 ; Patent No. 5780263
 ; GENERAL INFORMATION:
 ; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
 ; TITLE OF INVENTION: Human CCN-Like Growth Factor
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
 ; STREET: 6 BECKER FARM ROAD
 ; CITY: ROSELAND
 ; STATE: NEW JERSEY
 ; COUNTRY: USA
 ; ZIP: 07068
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 INCH DISKETTE
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WORD PERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/468,847B
 ; FILING DATE: 6 June 1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33, 073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-468-847B-13

Query Match 80.1%; Score 1695; DB 1; Length 375;
Best Local Similarity 81.0%; Pred. No. 4.2e-140;
Matches 311; Conservative 20; Mismatches 41; Indels 12; Gaps 7;

QY 1 MSRIARALALVTLHLTLRLAL-STCPACHCPLEAPKCAPGVGLVBDGGCCCKVCAKQ 59
DB 1 MGSAGARP-ALAAALICLARIALGSPCAVCCCPAAAPQCAPGVGLVBDGGCCCKVCAKQ 59
QY 60 LNEDESKTOPCDHRTKGLCECNFGASSTALKGICRAQSEGRPCENSRITYONGESFQPNCOH 119
DB 60 LNEDESKTOPCDHRTKGLCECNFGASSTALKGICRAQSEGRPCENSRITYONGESFQPNCKH 119
QY 120 QCTCTIDGAVGCIPLCPQELSLPNLGCPNRLVYKVTGCCCEWVCDEDSIKDPMEDQGL 179
DB 120 QCTCTIDGAVGCIPLCPQELSLPNLGCPNRLVYKVTGCCCEWVCDEDSIKDPMEDQGL 179
QY 180 GKELGFASVELELRNNELIIVAGKGRSLKRLPVGMEPR--RLVYVLPDGGKCIYOTTSMS 237
DB 178 SKREGLDASEELRNELIIVAGKGRSLKRLPVGMEPR--RLVYVLPDGGKCIYOTTSMS 232
QY 238 QCSKTGCTGISTRTVNDNPECRVYKTRICEVRPCGQVYSSLSKKKKSKTKKSPPEVR 297
DB 233 QCSKTGCTGISTRTVNDNPECRVYKTRICEVRPCGQVYSSLSKKKKSKTKKSPPEVR 292
QY 298 FTYAGCLSVKRYRKYRGCSYVDGRCCTPQTLRTYKMRRCDEDEFTSKNVMIIOSCKNY 357
DB 293 FTYAGCSVKKRYRKYRGCSYVDGRCCTPQTLRTYKMRRCDEDEFTSKNVMIIOSCKNY 352
QY 358 NCPHANEAPFPYRLFENDIHKFRD 381
DB 353 NCPHANEAPFPYRLVNDIHKFRD 375

RESULT 8
US-09-582-337-2
Sequence 2, Application US/09582337
Patent No. 6562618

GENERAL INFORMATION:
APPLICANT: Japan Tobacco, Inc.
TITLE OF INVENTION: Monoclonal Antibody Against Connective Tissue Growth Factor
TITLE OF INVENTION: and Medicinal Uses Thereof
FILE REFERENCE: JI-0093CT
CURRENT APPLICATION NUMBER: US/09/582,337
CURRENT FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: JP P1997-367699
PRIOR FILING DATE: 1997-12-25
PRIOR APPLICATION NUMBER: JP P1998-356183
PRIOR FILING DATE: 1998-12-15
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2
LENGTH: 347
TYPE: PRT
ORGANISM: Rat
US-09-582-337-2

Query Match 45.0%; Score 953; DB 4; Length 347;
Best Local Similarity 46.2%; Pred. No. 2.2e-75;

Matches 176; Conservative 58; Mismatches 107; Indels 40; Gaps 7;

QY 1 MSRIARALALVTLHLTLRLAL-STCPACHCPLEAPKCAPGVGLVBDGGCCCKVCAK 58
DB 1 MGSAGARP-ALAAALICLARIALGSPCAVCCCPAAAPQCAPGVGLVBDGGCCCKVCAK 60
QY 59 QLNEDSKTOPCDHRTKGLCECNFGASSTALKGICRAQSEGRPCENSRITYONGESFQPNCOH 118
DB 61 QLNEDSKTOPCDHRTKGLCECNFGASSTALKGICRAQSEGRPCENSRITYONGESFQPNCKH 119
QY 119 HOCTCTIDGAVGCIPLCPQELSLPNLGCPNRLVYKVTGCCCEWVCDEDSIKDPMEDQGL 178
DB 120 HOCTCTIDGAVGCIPLCPQELSLPNLGCPNRLVYKVTGCCCEWVCDEDSIKDPMEDQGL 167
QY 179 GKELGFASVELELRNNELIIVAGKGRSLKRLPVGMEPR--RLVYVLPDGGKCIYOTTSMS 237
DB 168 GKELGFASVELELRNNELIIVAGKGRSLKRLPVGMEPR--RLVYVLPDGGKCIYOTTSMS 205
QY 238 QCSKTGCTGISTRTVNDNPECRVYKTRICEVRPCGQVYSSLSKKKKSKTKKSPPEVR 297
DB 206 ACSKTGCTGISTRTVNDNPECRVYKTRICEVRPCGQVYSSLSKKKKSKTKKSPPEVR 265
QY 298 FTYAGCLSVKRYRKYRGCSYVDGRCCTPQTLRTYKMRRCDEDEFTSKNVMIIOSCKNY 357
DB 266 FTYAGCSVKKRYRKYRGCSYVDGRCCTPQTLRTYKMRRCDEDEFTSKNVMIIOSCKNY 325
QY 358 NCPHANEAPFPYRLFENDIHKFRD 376
DB 326 NCPHANEAPFPYRLVNDIHKFRD 346

RESULT 9
US-09-292-036-3
Sequence 3, Application US/09292036
Patent No. 6358741

GENERAL INFORMATION:
APPLICANT: FIBROGEN, INC.
APPLICANT: SCHMIDT, Brian
APPLICANT: ALLEN, Margaret
APPLICANT: SVERDRUP, Fran
APPLICANT: CARMICHAEL, David
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR (CTGF) AND METHODS OF USE
FILE REFERENCE: FIBRO100-1
CURRENT APPLICATION NUMBER: US/09/292,036
CURRENT FILING DATE: 1999-04-14
PRIOR APPLICATION NUMBER: US 09/292,036
PRIOR FILING DATE: 1999-04-14
PRIOR APPLICATION NUMBER: US 09/187,478
PRIOR FILING DATE: 1998-11-06
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 348
TYPE: PRT
ORGANISM: Mouse
US-09-292-036-3

Query Match 45.0%; Score 952.5; DB 4; Length 348;
Best Local Similarity 46.6%; Pred. No. 2.5e-75;
Matches 178; Conservative 57; Mismatches 106; Indels 41; Gaps 8;

QY 1 MSRIARALALVTLHLTLRLAL-STCPACHCPLEAPKCAPGVGLVBDGGCCCKVCAK 57
DB 1 MGSAGARP-ALAAALICLARIALGSPCAVCCCPAAAPQCAPGVGLVBDGGCCCKVCAK 60
QY 58 QLNEDSKTOPCDHRTKGLCECNFGASSTALKGICRAQSEGRPCENSRITYONGESFQPNCOH 117
DB 61 QLNEDSKTOPCDHRTKGLCECNFGASSTALKGICRAQSEGRPCENSRITYONGESFQPNCKH 119
QY 118 HOCTCTIDGAVGCIPLCPQELSLPNLGCPNRLVYKVTGCCCEWVCDEDSIKDPMEDQGL 177
DB 120 HOCTCTIDGAVGCIPLCPQELSLPNLGCPNRLVYKVTGCCCEWVCDEDSIKDPMEDQGL 168

QY 178 LKGLKELGPASVEVELRNNEELAVGGRSLKRL-PVFGMEPRLLYNPLOGOKCIYOTTSW 236
Db 169 -----KDRTAVGPAALAYRLIEDTFGGDPPTMM-----RANCLVOTTEW 205
QY 237 SOCSKTCGTGISTRTVNDNPECLVETRICERPCGQPVYSSLKGGKSKTKRSKPEPY 296
Db 206 SACSCTGCGISRTVNDNTPFCHLEKOSRLCWRPCPADLEENIKKKKCIKRPKIAKRV 265
QY 297 RFTYAGCLSVKKYRPYCGSCVDGRCCTPOLRTVYMRPRCEDGFEFFSKNNVMIOQCKCN 356
Db 266 KFELSGCTSVKTYRAKFCVCTDGRCTPHRTTLVEFKCPDGEIMKKNNMFIKTCACH 325
QY 357 YNCPHANEAFPEY--RLFNDI 376
Db 326 YNCPGNDIPESLYRKMYGDM 347

RESULT 10

US-08-468-847B-15
Sequence 15, Application US/08468847B
Patent No. 5780263
GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARBILA, BYRNE, BAIN, GILFILLIAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,847B
FILING DATE: 6 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-468-847B-15

Query Match 44.8%; Score 948.5; DB 1; Length 348;
Best Local Similarity 46.3%; Pred. No. 5.5e-75;
Matches 177; Conservative 58; Mismatches 106; Indels 41; Gaps 8;
QY 1 MSSRIARALAVVTLHL-TRLAL-STCPAACHPLE-APKCAPGVGVRODGGCGCKYCA 57
Db 1 MLASVAGPISLALVLLALCTRPATGQDCSAQCQCAAEAPHPCPAGVSLVLDGCGCRVCA 60
QY 58 KQLNDCSKTOPCDHTKGLCNFGASSTALKGICRAQSEGRPCCEYNRIYQNGESFQPN 117
Db 61 KQLGELCTERBDCPHKGLGDFGSPANRKKIGVCIAR-DGAPCVFGGSYVRSGESFQSSC 119

QY 118 QHOCTCIDGAVGCIPLCPQELSLPNLGCENPRLVKTGQCCSEWVEDSDSIKDPMEDQGS 177
Db 120 KYQCTCIDGAVGCVPLCSMDVRLBSPDCEFPFRVKKLPBGCKKMWCDPE-----168
QY 178 LKGLKELGPASVEVELRNNEELAVGGRSLKRL-PVFGMEPRLLYNPLOGOKCIYOTTSW 236
Db 169 -----KDRTAVGPAALAYRLIEDTFGGDPPTMM-----RANCLVOTTEW 205
QY 237 SOCSKTCGTGISTRTVNDNPECLVETRICERPCGQPVYSSLKGGKSKTKRSKPEPY 296
Db 206 SACSCTGCGISRTVNDNTPFCHLEKOSRLCWRPCPADLEENIKKKKCIKRPKIAKRV 265
QY 297 RFTYAGCLSVKKYRPYCGSCVDGRCCTPOLRTVYMRPRCEDGFEFFSKNNVMIOQCKCN 356
Db 266 KFELSGCTSVKTYRAKFCVCTDGRCTPHRTTLVEFKCPDGEIMKKNNMFIKTCACH 325
QY 357 YNCPHANEAFPEY--RLFNDI 376
Db 326 YNCPGNDIPESLYRKMYGDM 347

RESULT 11

US-09-142-569-6
Sequence 6, Application US/09142569
Patent No. 6413735
GENERAL INFORMATION:
APPLICANT: Lau, Lester F.
TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/142,569
FILING DATE: 02-Apr-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28758/33766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: "Fisp12 amino acid sequence"
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-142-569-6

Query Match 44.8%; Score 948.5; DB 4; Length 348;
Best Local Similarity 46.3%; Pred. No. 5.5e-75;
Matches 177; Conservative 58; Mismatches 106; Indels 41; Gaps 8;
QY 1 MSSRIARALAVVTLHL-TRLAL-STCPAACHPLE-APKCAPGVGVRODGGCGCKYCA 57
Db 1 MLASVAGPISLALVLLALCTRPATGQDCSAQCQCAAEAPHPCPAGVSLVLDGCGCRVCA 60

```

QY 58 KOLNEOSKTOPCDHTKGLSENFNGASSTALICGAOSEBGPCEVNSIYONGESPQNC 117
Dd 61 KQJGEICTEDPDCDDPHKGLFCDFGSPANKRGVCTAK-DGAPCVFGSGVSRSGESFQSC 119
QY 118 QHOCTCIDGAVGICPPLCPOELSLPNIGCPNPLVYVTOCCCEEWVCDEDSIKDMEODG 177
Dd 120 KYOCTCLDGAVGCVPLCSMDVRLPSPDCDFPRKRVKLPCKCKEWCDEP----- 168
QY 178 LLGKELGPDASEVELTRNNELLVAGKGSGLKRL-PVFGMEPRILYNPLQOGKCIYQTTW 236
Dd 169 -----KDRTAVGALAAAYLLEDTFEBDPDTM-----RANCLVQTTW 205
QY 237 SQCSKTGCGISTRTVNDNPNBECRLVKEPTRICEVPRQGPVYSLSLKGCCKSCTKKSPEV 296
Dd 206 SACSCTCGMGISTRTVNDNTPCRLKQSLCVMRCEADLEBNIKKGRCTIRPFIKAPV 265
QY 297 RFTYAGCLSYKKTRPVYCGSCYVDGRCCTPOLTRTYKMFREDEGETSKNVMQOSCKN 356
Dd 266 KPELSCTSKTYIRAFKGCVCYDGRCTPHRTTLLPVEFKCPDEIMKKMMFIKTCAH 325
QY 357 YNCPHANEAAFPY--RLFNDI 376
Dd 326 YNCPGNDNDFESLIYRKMTGDM 347

```

```

RESULT 12 478-2
US-09-187-478-2
: Sequence 2, Application US/09187478
: Patent No. 6348329
: GENERAL INFORMATION:
: APPLICANT: Schmidt, Brian F.
: APPLICANT: Allen, Margaret L.
: TITLE OF INVENTION: Connective Tissue Growth (CTGF) And Methods Of Use
: FILE REFERENCE: 08766/004001
: CURRENT APPLICATION NUMBER: US/09/187,478
: CURRENT FILING DATE: 1998-11-06
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 347
: TYPE: PRT
: ORGANISM: NO. 6348329ma1 Rate Kidney Fibroblast
US-09-187-478-2

```

Query Match	44.88;	Score 948;	DB 4;	Length 347;
Best Local Similarity	46.38;	Pred. No. 6e-75;		
Matches 177;	Conservative 58;	Mismatches 105;	Indels 42;	Gaps 9

QY MSRRIRALAIYVTLITLITRL--STCPAACHPLE-APKCAPVGIYRGCGCKYCAK 58
 Db 1 MASVAGPYSIALYLLITLITRPAITGDDCSAGCCAAEAPAPAGSLVLDGGCKRYCAK 60

QY QLNEDCSKTPQCDHTKGLKCNFGAASSTALKGICRAQSGRCEYNRIYONGESPOQC 118
 Db 61 QLQELCTIERNPCDPHSLIFCDPFGSPANKRIIGVCTAK-DGACICVGVGSSYSGESFOSSCK 119

QY HOGCTIDGAVGCIPLCPQELSLPNIGCENPNRYKTKGQCCGEWCDDESLKDPMEQDGL 178
 Db 120 YQCTCIDGAVGCVPLCSMDYRLPSPDCCFPFRVYKILPGKCCBWCDEP----- 167

QY 179 LKELCFDASEVELTRNNELIAYGKGRSLKRL-PYEGMEPRILYNPLQOGCKCIYOTWS 237
 Db 168 -----KDRTPVYGALAAAYRLIEDTFEGDPTMM-----RANCIYOTWEWS 205

QY QCSKTCGTGISTRVINDNEPCLRYEETFLICEYRPOGQVYSLKKGKCKSTKSPSPR 297
 Db 206 ACGSKTCGMGISTRVINDNTFCRLERQIILCWVRPEADLLENIKKGKCIIRTPRIAPYK 265

QY 298 FYTAGLSLKYKTRPKYCGSCVDGRCCQTQOLIRTYKAMRRCDGETFSKNVMTIOSCKNY 357
 Db 266 FELSGTSVAKYTRAKFCVGDGRCCQTCHARTTTLPEVERCKDDGIMKNMFIKTCACH 325

Qy 358 NCPHANEAFP--FYR-LENDI 376
||| | : || :|| :: | :
Db 326 NCPGDND-IFPCMYRKMVGDM 346

RESULT 13
 US-09-292-036-2
 ; Sequence 2, Application US/09292036
 ; Patent No. 6358741
 ; GENERAL INFORMATION:
 ; APPLICANT: FIBROGEN, INC
 ; APPLICANT: SCHMIDT, Brian
 ; APPLICANT: ALLEN, Margaret
 ; APPLICANT: SVERDRUP, Fran
 ; APPLICANT: CAMICHAEL, David
 ; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR (CTGF) AND METHODS OF USE
 ; FILE REFERENCE: FIBROL100-1
 ; CURRENT APPLICATION NUMBER: US/09/292,036
 ; CURRENT FILING DATE: 1999-04-14
 ; PRIOR APPLICATION NUMBER: US 09/292,036
 ; PRIOR FILING DATE: 1999-04-14
 ; PRIOR APPLICATION NUMBER: US 09/187,478
 ; PRIOR FILING DATE: 1998-11-06
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 347
 ; TYPE: PRT
 ; ORGANISM: Rat
 ; US-09-292-036-2

Query Match	44.8%;	Score 948;	DB 4;	Length 347;
Best Local Similarity	46.3%;	Pred. No. 6e-75;		
Matches 177;	Conservative 58;	Mismatches 105;	Indels 42;	Gaps 9;

```

OY      1  MSSRIARLALVLTLLHLTRLAL-STCPACACPLE-APRCAGVGLVPRGCCCKYAK  58
Db      1  MIASVAGVSTALVLTCTTRPATGTGDCSQCCQAFAARCPAGVSLVTDGCCCKRYAK  60
OY      59  QLNEDCSKTOPCDHTKGLCECNFGASSTALKGICRAQSEGRPCENRYONGSEFQPMQC  118
Db      61  QUGELCTERDPDCKHSLFCDFGRSPANRIGVCTAK-DGAPCVGGSVYRSRGSFSQCK  119
OY      119  HOCTCIDAGVCTIPLCPQELSLPNIGCPNRLVKYTGCCCEBWCDEDSIKDPMEDDGL  178
Db      120  YQCTCIDAGVGVCPICSDMYRLPSPDCPPFRVKYLGKCCQEWVCDEP-----  167
OY      179  LKGLGFGDASVEELRNNELIIVAGKRSJLKL-PVNGMEPRILIYNLQOGKIVOTTSWS  237
Db      168  -----KDRIVVGPALAVYRLIEDTFGGPDPTMM-----RANCLIVOTTEWS  205
OY      238  QCSKTGCTGISTRTVNDNDECELRVETRIACEVRPGQAPYSSLSLKGGKCSKTKKSEPYR  297
Db      206  ACSKTCGNGISIRRVYNDMTFECRLFNQIRCMWRPCPADLEENIKKKCKCIRFPKIAKPYR  265
OY      298  FTVAGCLISVKKIRPYCGSCVDGRCCTPOLNFTVKAIRPCDEGEYFSKVNMMIQQSKCNY  357
Db      266  FELSCGTSVKTYRAAFKCGCTGDCRCCTPHRTYTLVPEFKCPDGEIIMKKNMMEFKTCAHY  325
OY      358  NCPHANEAPF--EYR-LENDI  376
Db      326  NCPGNDN-IFPCMYIRKMYIGDM  346

```

RESULT 14
US-08-167-628-2
Sequence: 2, Application US/08167628
Patent No. 5408040
GENERAL INFORMATION:
APPLICANT: Grotenordt, Gary R.
APPLICANT: Brahnam Jr., Douglas M.,
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
NUMBER OF SEQUENCES: 2

```

CORRESPONDENCE ADDRESS:
ADDRESS: Spensley Horn Judas & Lubitz
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/167,628
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/752,427
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr. Ph.D., John W.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: PD-1294
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-455-5100
TELEFAX: 619-455-5110
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 349 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-167-628-2

Query Match
Best Local Similarity 46.8%; Score 947; DB 1; Length 349;
Matches 174; Conservative 58; Mismatches 101; Indels 40; Gaps 8;

QY 9 LALVVTLLHLRLAL-STCPAACHCPL-APRCAPGVGLVRDGGCCCKVCAKOLNEDCSK 66
DB 11 VAFVLLALCSRPAGVQNSGPCRCPDEPAPRCAGVSLVLDGCCRCVCAKOLGELCTE 70
QY 67 TOPCHTGLNEFGASSFALKGICRAOSEGRPCENSRITYNGSFOPNCHOCCTIDG 126
DB 71 RDCPDHKGFLFCDFSPANKRGVCTAK-DGAPCIJFGTVYRSGSFQSSCKYQCTCLDG 129
QY 127 AVGCIPLPQELSLNLGCPNRLVKTGOCCEEWCDSDSIKDPMEDQDGLGKELGFD 186
DB 130 AVGCMPLCSMDVRLSPDCPFPRRVKRLPGKCEEWCDSDSIKDPMEDQDGLGKELGFD 172
QY 187 ASEVELTRNNELIAVKGSRSLKL-PVFGMEPRILYNPLQOGKCIYVOTTSMSQSKTGT 245
DB 173 -----TVGPAALAAVRLDFTGPDPTMI-----RANCLVQTTMWSAGSKTCGM 215
QY 246 GISTRTVNDNPECRVYKERICVEVRPCQGPVYSSLSKKGKSKTKSPVAFYTAGCLS 305
DB 216 GISTRTVNDNASCRIEKSRLCMVRPCEADLEENIKKSKCIRTPIKISPIKIFELSGCTS 275
QY 306 VKYRPKYGSCVDCRCCTPOLRTVKKMRFCEDGETFSKNVMYIOSCKNYNCPHANEA 365
DB 276 MKTYRAKFCGYCTDRCTPHRTTLLPYEFKCPDGEVAKKMMFIKTCAHINCAGDNDI 335
QY 366 AFPEY--RLFNDI 376
DB 336 FESLYRRKMYGDM 348
```

```

APPLICANT: Bradham Jr., Douglas M.,
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Spensley Horn Judas & Lubitz
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/386,680
FILING DATE: 10-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/167,628
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr. Ph.D., John W.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: PD-1294
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-455-5100
TELEFAX: 619-455-5110
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 349 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-386-680-2

Query Match
Best Local Similarity 44.8%; Score 947; DB 1; Length 349;
Matches 174; Conservative 58; Mismatches 101; Indels 40; Gaps 8;

QY 9 LALVVTLLHLRLAL-STCPAACHCPL-APRCAPGVGLVRDGGCCCKVCAKOLNEDCSK 66
DB 11 VAFVLLALCSRPAGVQNSGPCRCPDEPAPRCAGVSLVLDGCCRCVCAKOLGELCTE 70
QY 67 TOPCHTGLNEFGASSFALKGICRAOSEGRPCENSRITYNGSFOPNCHOCCTIDG 126
DB 71 RDCPDHKGFLFCDFSPANKRGVCTAK-DGAPCIJFGTVYRSGSFQSSCKYQCTCLDG 129
QY 127 AVGCIPLPQELSLNLGCPNRLVKTGOCCEEWCDSDSIKDPMEDQDGLGKELGFD 186
DB 130 AVGCMPLCSMDVRLSPDCPFPRRVKRLPGKCEEWCDSDSIKDPMEDQDGLGKELGFD 172
QY 187 ASEVELTRNNELIAVKGSRSLKL-PVFGMEPRILYNPLQOGKCIYVOTTSMSQSKTGT 245
DB 173 -----TVGPAALAAVRLDFTGPDPTMI-----RANCLVQTTMWSAGSKTCGM 215
QY 246 GISTRTVNDNPECRVYKERICVEVRPCQGPVYSSLSKKGKSKTKSPVAFYTAGCLS 305
DB 216 GISTRTVNDNASCRIEKSRLCMVRPCEADLEENIKKSKCIRTPIKISPIKIFELSGCTS 275
QY 306 VKYRPKYGSCVDCRCCTPOLRTVKKMRFCEDGETFSKNVMYIOSCKNYNCPHANEA 365
DB 276 MKTYRAKFCGYCTDRCTPHRTTLLPYEFKCPDGEVAKKMMFIKTCAHINCAGDNDI 335
QY 366 AFPEY--RLFNDI 376
DB 336 FESLYRRKMYGDM 348
```

Search completed: August 5, 2003, 14:08:02

Wed Aug 6 07:53:12 2003

Job time : 15.0368 secs

us-09-495-448a-4.rai

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:06:57; Search time 34.5908 Seconds

(Without alignments)
1308.079 Million cell updates/sec

Title: US-09-495-448a-4

Perfect score: 2116
Sequence: 1 MSSRIARALALVLTILHLTR.....ANERAPFFYRLPNDIHKFRD 381

Scoring table:

BLOSUM62
Gapop 10.0, Gapept 0.5

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 451899

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubppaa/PCR_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubppaa/PCRUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep.*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2116	100.0	381	14	US-10-053-753-4
2	2106	99.5	381	11	US-09-901-910-2
3	2106	99.5	381	15	US-10-294-796-2
4	2106	99.5	455	9	US-09-925-101-1432
5	2098	99.5	381	15	US-10-205-823-84
6	1971.5	93.2	374	9	US-09-853-6258-12
7	1971.5	93.2	375	11	US-09-901-910-7
8	1929	91.2	379	14	US-09-853-6258-11
9	1929	91.2	379	14	US-10-053-753-2
10	1695	80.1	375	9	US-09-853-6258-13
11	955	45.1	347	15	US-10-245-977-7
12	952.5	45.0	348	14	US-10-101-040-3
13	948.5	44.8	348	9	US-09-853-6258-15
14	948.5	44.8	348	14	US-10-053-753-6
15	948.5	44.8	348	15	US-10-245-977-8

16	948	44.8	347	14	US-10-080-173-2	Sequence 2, Appl1
17	948	44.8	347	14	US-10-101-040-2	Sequence 2, Appl1
18	947	44.8	349	14	US-10-101-040-4	Sequence 4, Appl1
19	947	44.8	349	14	US-10-011-859-26	Sequence 26, Appl1
20	947	44.8	349	14	US-10-053-753-8	Sequence 8, Appl1
21	947	44.8	349	15	US-10-060-036-173	Sequence 173, Appl
22	947	44.8	349	15	US-10-171-111-46	Sequence 46, Appl
23	947	44.8	349	15	US-10-205-823-78	Sequence 78, Appl
24	947	44.8	349	15	US-10-245-977-2	Sequence 2, Appl1
25	938.5	44.4	348	9	US-09-853-6258-14	Sequence 14, Appl1
26	933	44.1	349	15	US-10-245-977-5	Sequence 5, Appl1
27	906	42.8	349	15	US-10-245-977-6	Sequence 6, Appl1
28	851.5	40.2	351	9	US-09-853-6258-16	Sequence 16, Appl
29	827.5	39.1	357	9	US-09-853-6258-17	Sequence 17, Appl
30	827.5	39.1	357	14	US-10-011-859-25	Sequence 25, Appl
31	767	36.2	367	14	US-10-001-054-50	Sequence 50, Appl
32	767	36.2	367	15	US-10-112-267-4	Sequence 4, Appl1
33	767	36.2	367	15	US-10-112-267-8	Sequence 8, Appl1
34	766	36.2	367	15	US-10-112-267-7	Sequence 7, Appl1
35	766	36.2	367	15	US-10-112-267-22	Sequence 22, Appl
36	760.5	35.9	345	15	US-10-112-267-3	Sequence 3, Appl1
37	760.5	35.9	345	15	US-10-112-267-6	Sequence 6, Appl1
38	759.5	35.9	345	15	US-10-112-267-5	Sequence 5, Appl1
39	759.5	35.9	345	15	US-10-112-267-21	Sequence 21, Appl
40	758	35.8	367	15	US-10-112-267-12	Sequence 12, Appl
41	754.5	35.7	345	15	US-10-112-267-11	Sequence 11, Appl
42	613	29.0	339	15	US-10-112-267-36	Sequence 36, Appl
43	613	29.0	354	14	US-10-011-859-2	Sequence 2, Appl1
44	613	29.0	354	15	US-10-112-267-37	Sequence 37, Appl
45	612	28.9	339	15	US-10-112-267-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-10-053-753-4
Sequence 4, Application US/10053753
Publication No. US20020150986A1

GENERAL INFORMATION:
APPLICANT: Lau, Lester F.
TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/053.753
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28758/33766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 381 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: "Human Cyt61 amino acid sequence"
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-053-753-4

Query Match 100.0%; Score 2116; DB 14; Length 381;
Best Local Similarity 100.0%; Pred. No. 2,9e-173;
Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSRIARALAVVTLHLTRLALSTCPACCHCLEAPKCAPGVGLVDDGGCCCKVCAKOL 60
DB 1 MSSRIARALAVVTLHLTRLALSTCPACCHCLEAPKCAPGVGLVDDGGCCCKVCAKOL 60
QY 61 NEDCSKTQPCDHTKGLKGCNFGASSTALKGICRAQSGRPECEYNSRIYONGESFQPCQHQ 120
DB 61 NEDCSKTQPCDHTKGLKGCNFGASSTALKGICRAQSGRPECEYNSRIYONGESFQPCQHQ 120
QY 121 CTCIDAVAGCIPCLPQELSLPNIKCPNRLVKTGGCCCEWVDEDSIKDPMEDDQGLLG 180
DB 121 CTCIDAVAGCIPCLPQELSLPNIKCPNRLVKTGGCCCEWVDEDSIKDPMEDDQGLLG 180
QY 181 KELGPDASEVELTRNNELIYAVGKSLKRLPVFGMEPRILYNPLOGOKCIYVTTSMSCS 240
DB 181 KELGPDASEVELTRNNELIYAVGKSLKRLPVFGMEPRILYNPLOGOKCIYVTTSMSCS 240
QY 241 KTCGTGISTRVINDNECHRLVETRIICEVRPGQOPYSSILKKGKCKSKTKKSPPEVFTY 300
DB 241 KTCGTGISTRVINDNECHRLVETRIICEVRPGQOPYSSILKKGKCKSKTKKSPPEVFTY 300
QY 301 AGCLSVKRYKPKYKCGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIQQCKVNCNP 360
DB 301 AGCLSVKRYKPKYKCGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIQQCKVNCNP 360
QY 361 HANEAAFPFRLFNDIHKFRD 381
DB 361 HANEAAFPFRLFNDIHKFRD 381

RESULT 2

US-09-901-910-2
; Sequence 2, Application US/09901910
; Publication No. US20030012768A1
; GENERAL INFORMATION:
; APPLICANT: Li, Haodong
; APPLICANT: Adams, Mark
; APPLICANT: Calenda Valerie
; TITLE OF INVENTION: Connective Tissue Growth Factor-2
; FILE REFERENCE: P126P2
; CURRENT APPLICATION NUMBER: US/09/901,910
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/348,815
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 08/459,101
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: PCT/US94/07736
; PRIOR FILING DATE: 1994-07-12
; PRIOR APPLICATION NUMBER: 60/217,402
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/291,642
; PRIOR FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; LENGTH: 381
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-901-910-2

Query Match 99.5%; Score 2106; DB 11; Length 381;
Best Local Similarity 99.5%; Pred. No. 2,1e-172;
Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALAVVTLHLTRLALSTCPACCHCLEAPKCAPGVGLVDDGGCCCKVCAKOL 60
DB 1 MSSRIARALAVVTLHLTRLALSTCPACCHCLEAPKCAPGVGLVDDGGCCCKVCAKOL 60
QY 61 NEDCSKTQPCDHTKGLKGCNFGASSTALKGICRAQSGRPECEYNSRIYONGESFQPCQHQ 120
DB 61 NEDCSKTQPCDHTKGLKGCNFGASSTALKGICRAQSGRPECEYNSRIYONGESFQPCQHQ 120
QY 121 CTCIDAVAGCIPCLPQELSLPNIKCPNRLVKTGGCCCEWVDEDSIKDPMEDDQGLLG 180
DB 121 CTCIDAVAGCIPCLPQELSLPNIKCPNRLVKTGGCCCEWVDEDSIKDPMEDDQGLLG 180
QY 181 KELGPDASEVELTRNNELIYAVGKSLKRLPVFGMEPRILYNPLOGOKCIYVTTSMSCS 240
DB 181 KELGPDASEVELTRNNELIYAVGKSLKRLPVFGMEPRILYNPLOGOKCIYVTTSMSCS 240
QY 241 KTCGTGISTRVINDNECHRLVETRIICEVRPGQOPYSSILKKGKCKSKTKKSPPEVFTY 300
DB 241 KTCGTGISTRVINDNECHRLVETRIICEVRPGQOPYSSILKKGKCKSKTKKSPPEVFTY 300
QY 301 AGCLSVKRYKPKYKCGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIQQCKVNCNP 360
DB 301 AGCLSVKRYKPKYKCGSCVDGRCTPOLTRTVKMRFCEDGETFSKNVMMIQQCKVNCNP 360
QY 361 HANEAAFPFRLFNDIHKFRD 381
DB 361 HANEAAFPFRLFNDIHKFRD 381

RESULT 3

US-10-294-796-2
; Sequence 2, Application US/10294796
; Publication No. US20030078391A1
; GENERAL INFORMATION:
; APPLICANT: Li, Haodong et al.
; TITLE OF INVENTION: Connective Tissue Growth Factor-2
; FILE REFERENCE: P126PID2
; CURRENT APPLICATION NUMBER: US/10/294,796
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 09/348,815
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US 08/459,101
; PRIOR FILING DATE: 1995-06-02
; PRIOR APPLICATION NUMBER: PCT/US94/07736
; PRIOR FILING DATE: 1994-07-12
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-294-796-2

Query Match 99.5%; Score 2106; DB 15; Length 381;
Best Local Similarity 99.5%; Pred. No. 2,1e-172;
Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALAVVTLHLTRLALSTCPACCHCLEAPKCAPGVGLVDDGGCCCKVCAKOL 60
DB 1 MSSRIARALAVVTLHLTRLALSTCPACCHCLEAPKCAPGVGLVDDGGCCCKVCAKOL 60
QY 61 NEDCSKTQPCDHTKGLKGCNFGASSTALKGICRAQSGRPECEYNSRIYONGESFQPCQHQ 120
DB 61 NEDCSKTQPCDHTKGLKGCNFGASSTALKGICRAQSGRPECEYNSRIYONGESFQPCQHQ 120
QY 121 CTCIDAVAGCIPCLPQELSLPNIKCPNRLVKTGGCCCEWVDEDSIKDPMEDDQGLLG 180
DB 121 CTCIDAVAGCIPCLPQELSLPNIKCPNRLVKTGGCCCEWVDEDSIKDPMEDDQGLLG 180
QY 181 KELGPDASEVELTRNNELIYAVGKSLKRLPVFGMEPRILYNPLOGOKCIYVTTSMSCS 240
DB 181 KELGPDASEVELTRNNELIYAVGKSLKRLPVFGMEPRILYNPLOGOKCIYVTTSMSCS 240

QY 241 KTCGTGISTRTVNDNPECRIVKTRICEVRCGOPYVSSLKGGKCKSKTKKSPPEVFTY 300
Db 241 KTCGTGISTRTVNDNPECRIVKTRICEVRCGOPYVSSLKGGKCKSKTKKSPPEVFTY 300
QY 301 AGCISVKKYKRPKYCGSCVDGRCTPOLTRTVKMRFRCEDEFTSKNMMIOSCKCNYNCP 360
Db 301 AGCISVKKYKRPKYCGSCVDGRCTPOLTRTVKMRFRCEDEFTSKNMMIOSCKCNYNCP 360
QY 361 HANEAFPEYRLFNDDIHKFRD 381
Db 361 HANEAFPEYRLFNDDIHKFRD 381

RESULT 4
US-09-925-301-1432
; Sequence 1432, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P1106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/055882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 1432
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-1432

Query Match 99.5%; Score 2106; DB 9; Length 455;
Best Local Similarity 99.5%; Pred. No. 2.5e-172;
Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSRIARALALVYTLHLTRIALSTCPACHCPLEAPKCAPGVGLVBDGGCCCKVCAKOL 60
Db 75 MSSRIARALALVYTLHLTRIALSTCPACHCPLEAPKCAPGVGLVBDGGCCCKVCAKOL 134
QY 61 NEDCSKTOPCDHRTGELCNFGASSTALKGICRAOSEBPEYNSRITQONESPOPNCKHQ 120
Db 135 NEDCSKTOPCDHRTGELCNFGASSTALKGICRAOSEBPEYNSRITQONESPOPNCKHQ 194
QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPRLVKTGQCCCEMWVCEDESIKPMEDQDGLG 180
Db 195 CTCIDGAVGCIPLCPQELSLPNLGCNPRLVKTGQCCCEMWVCEDESIKPMEDQDGLG 254
QY 181 KELGFDASEVELTRNNELIANGKRSILKRLPVGMEBRILYNPLOGKCIYQTTSMGQS 240
Db 255 KELGFDASEVELTRNNELIANGKRSILKRLPVGMEBRILYNPLOGKCIYQTTSMGQS 314
QY 241 KTCGTGISTRTVNDNPECRIVKTRICEVRCGOPYVSSLKGGKCKSKTKKSPPEVFTY 300
Db 315 KTCGTGISTRTVNDNPECRIVKTRICEVRCGOPYVSSLKGGKCKSKTKKSPPEVFTY 374
QY 301 AGCISVKKYKRPKYCGSCVDGRCTPOLTRTVKMRFRCEDEFTSKNMMIOSCKCNYNCP 360
Db 375 AGCISVKKYKRPKYCGSCVDGRCTPOLTRTVKMRFRCEDEFTSKNMMIOSCKCNYNCP 434
QY 361 HANEAFPEYRLFNDDIHKFRD 381
Db 435 HANEAFPEYRLFNDDIHKFRD 455

RESULT 5
US-10-205-823-84
; Sequence 84, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:

APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Endege, Wilson O.
APPLICANT: Gannavarapu, Manjula
APPLICANT: Gorbatcheva, Bella
APPLICANT: Hoersch, Sebastian
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Monsey, Angela M.
APPLICANT: Glatz, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-84

Query Match 99.1%; Score 2098; DB 15; Length 381;
Best Local Similarity 99.0%; Pred. No. 1e-171;
Matches 377; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSSRIARALALVYTLHLTRIALSTCPACHCPLEAPKCAPGVGLVBDGGCCCKVCAKOL 60
Db 1 MSSRIARALALVYTLHLTRIALSTCPACHCPLEAPKCAPGVGLVBDGGCCCKVCAKOL 60
QY 61 NEDCSKTOPCDHRTGELCNFGASSTALKGICRAOSEBPEYNSRITQONESPOPNCKHQ 120
Db 61 NEDCSKTOPCDHRTGELCNFGASSTALKGICRAOSEBPEYNSRITQONESPOPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNLGCNPRLVKTGQCCCEMWVCEDESIKPMEDQDGLG 180
Db 121 CTCIDGAVGCIPLCPQELSLPNLGCNPRLVKTGQCCCEMWVCEDESIKPMEDQDGLG 180
QY 181 KELGFDASEVELTRNNELIANGKRSILKRLPVGMEBRILYNPLOGKCIYQTTSMGQS 240
Db 181 KELGFDASEVELTRNNELIANGKRSILKRLPVGMEBRILYNPLOGKCIYQTTSMGQS 240
QY 241 KTCGTGISTRTVNDNPECRIVKTRICEVRCGOPYVSSLKGGKCKSKTKKSPPEVFTY 300
Db 241 KTCGTGISTRTVNDNPECRIVKTRICEVRCGOPYVSSLKGGKCKSKTKKSPPEVFTY 300
QY 301 AGCISVKKYKRPKYCGSCVDGRCTPOLTRTVKMRFRCEDEFTSKNMMIOSCKCNYNCP 360
Db 301 AGCISVKKYKRPKYCGSCVDGRCTPOLTRTVKMRFRCEDEFTSKNMMIOSCKCNYNCP 360
QY 361 HANEAFPEYRLFNDDIHKFRD 381
Db 361 HANEAFPEYRLFNDDIHKFRD 381

RESULT 6
US-09-853-625B-12
; Sequence 12, Application US/09853625B
; Patent No. US20020049304A1
; GENERAL INFORMATION:

APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
CECCHI, STEWART & OLSTEIN

STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/853,625B

FILING DATE: 14-May-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/053,587

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.

REGISTRATION NUMBER: 33,073

REFERENCE/DOCKET NUMBER: 325800-442

TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:

LENGTH: 374 AMINO ACIDS

TYPE: AMINO ACID

STRANDEDNESS: <Unknown>

MOLECULE TYPE: PROTEIN

SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-09-853-625B-12

Query Match 93.2%; Score 1971.5; DB 9; Length 374;

Best Local Similarity 95.7%; Pred. No. 6,6e-161;

Matches 358; Conservative 2; Mismatches 13; Indels 1; Gaps 1;

1 MSSRIARALAVYTLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60

1 MSSRIARELAVYTLHLTRVGLSTCPADCHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60

61 NECCSTQPCDHTKGLGECNFGASSTALKGICRAQSGRCRCVNSRIYONGESFQPCQHQ 120

61 NECCSTQPCDHTKGLGECNFGASSTALKGICRAQSGRCRCVNSRIYONGESFQPCQHQ 120

121 CTCIDGAVG-CIPLCQELSLPMLGCPNRLVYVTCCEEWVCDSDSIKDPMEDDGL 179

121 CTCIDGAVG-CIPLCQELSLPMLGCPNRLVYVTCCEEWVCDSDSIKDPMEDDGL 179

121 CTCIDGAVG-CIPLCQELSLPMLGCPNRLVYVTCCEEWVCDSDSIKDPMEDDGL 179

180 GKEIGDASEVELTRNNELIANGKGRSLKRLPYFGMEPRILYNPLDGGKCIYQTTSMSC 239

180 GKEIGDASEVELTRNNELIANGKGRSLKRLPYFGMEPRILYNPLDGGKCIYQTTSMSC 239

181 GKEIGDASEVELTRNNELIANGKGRSLKRLPYFGMEPRILYNPLDGGKCIYQTTSMSC 240

240 SKTCGTGISTRYTNDNPECLVKEIRICEVRPGQGVYSSLKGGKCSKTKKSPPEVRFT 299

240 SKTCGTGISTRYTNDNPECLVKEIRICEVRPGQGVYSSLKGGKCSKTKKSPPEVRFT 299

300 YAGCLSVKRYRPYKCGSVDRGCTPQLTRVYKMRFRCDGEGFFSKNVMMIOSCKNYNC 359

300 YAGCLSVKRYRPYKCGSVDRGCTPQLTRVYKMRFRCDGEGFFSKNVMMIOSCKNYNC 359

301 YAGCLSVKRYRPYKCGSVDRGCTPQLTRVYKMRFRCDGEGFFSKNVMMIOSCKNYNC 360

360 PHANEAAPFYYRLF 373

361 PHANEAAPFYYRLF 374

RESULT 7
US-09-901-910-7
Sequence 7, Application US/09901910
Publication No. US20030012768A1

GENERAL INFORMATION:
APPLICANT: Li, Haodong

APPLICANT: Adams, Mark

TITLE OF INVENTION: Connective Tissue Growth Factor-2

FILE REFERENCE: P126P2

CURRENT APPLICATION NUMBER: US/09/901,910

CURRENT FILING DATE: 2001-07-11

PRIOR APPLICATION NUMBER: 09/348,815

PRIOR FILING DATE: 1999-07-08

PRIOR APPLICATION NUMBER: 08/459,101

PRIOR FILING DATE: 1995-06-02

PRIOR APPLICATION NUMBER: PCT/US94/07736

PRIOR FILING DATE: 1994-07-12

PRIOR APPLICATION NUMBER: 60/217,402

PRIOR FILING DATE: 2000-07-11

PRIOR APPLICATION NUMBER: 60/291,642

PRIOR FILING DATE: 2001-05-18

NUMBER OF SEQ ID NOS: 8

SOFTWARE: PatentIn version 3.0

SEQ ID NO: 7

LENGTH: 375

TYPE: PRT

ORGANISM: homo sapiens

US-09-901-910-7

Query Match 93.2%; Score 1971.5; DB 11; Length 375;

Best Local Similarity 95.7%; Pred. No. 6,7e-161;

Matches 358; Conservative 2; Mismatches 13; Indels 1; Gaps 1;

1 MSSRIARALAVYTLHLTRALSTCPAACHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60

1 MSSRIARELAVYTLHLTRVGLSTCPADCHCPLEAPKCAPGVGLVRDGGCCCKVCAKOL 60

61 NECCSTQPCDHTKGLGECNFGASSTALKGICRAQSGRCRCVNSRIYONGESFQPCQHQ 120

61 NECCSTQPCDHTKGLGECNFGASSTALKGICRAQSGRCRCVNSRIYONGESFQPCQHQ 120

121 CTCIDGAVG-CIPLCQELSLPMLGCPNRLVYVTCCEEWVCDSDSIKDPMEDDGL 179

121 CTCIDGAVG-CIPLCQELSLPMLGCPNRLVYVTCCEEWVCDSDSIKDPMEDDGL 179

121 CTCIDGAVG-CIPLCQELSLPMLGCPNRLVYVTCCEEWVCDSDSIKDPMEDDGL 179

180 GKEIGDASEVELTRNNELIANGKGRSLKRLPYFGMEPRILYNPLDGGKCIYQTTSMSC 239

180 GKEIGDASEVELTRNNELIANGKGRSLKRLPYFGMEPRILYNPLDGGKCIYQTTSMSC 239

181 GKEIGDASEVELTRNNELIANGKGRSLKRLPYFGMEPRILYNPLDGGKCIYQTTSMSC 240

240 SKTCGTGISTRYTNDNPECLVKEIRICEVRPGQGVYSSLKGGKCSKTKKSPPEVRFT 299

240 SKTCGTGISTRYTNDNPECLVKEIRICEVRPGQGVYSSLKGGKCSKTKKSPPEVRFT 299

300 YAGCLSVKRYRPYKCGSVDRGCTPQLTRVYKMRFRCDGEGFFSKNVMMIOSCKNYNC 359

300 YAGCLSVKRYRPYKCGSVDRGCTPQLTRVYKMRFRCDGEGFFSKNVMMIOSCKNYNC 359

301 YAGCLSVKRYRPYKCGSVDRGCTPQLTRVYKMRFRCDGEGFFSKNVMMIOSCKNYNC 360

360 PHANEAAPFYYRLF 373

361 PHANEAAPFYYRLF 374

RESULT 8

US-09-853-625B-11

Sequence 11, Application US/09853625B

Patent No. US20020049304A1

GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.

TITLE OF INVENTION: Human CCN-Like Growth Factor

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,

CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/853.625B
FILING DATE: 14-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/053.587
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: MULINS, J G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 379 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS: <Unknown>
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-853-625B-11

Query Match 91.2%; Score 1929; DB 9; Length 379;
Best Local Similarity 90.9%; Pred. No. 2.9e-157;
Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSSRIARALAVTLHLTRIALSTCPAACHCPLEAPKACAGVGLVRDGGCCCKVCAKOL 60
DB 1 MSSSTFRTLAVALVTLHLTRIALSTCPAACHCPLEAPKACAGVGLVRDGGCCCKVCAKOL 60
QY 61 NEDSKTOPCDHTGLEGCFNFGASSTALKGICRAOSEGRPCBYNSRITYONGESFOPNCKHQ 120
DB 61 NEDSKTOPCDHTGLEGCFNFGASSTALKGICRAOSEGRPCBYNSRITYONGESFOPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNLGCPNRLVKTGQCCCEWVCDSDSIKIDMEDODGLG 180
DB 121 CTCIDGAVGCIPLCPQELSLPNLGCPNRLVKTGQCCCEWVCDSDSIKIDMEDODGLG 180
QY 181 KELGFDASEVELTRNNELIAVGKRSILKRLPFVFGMEPRILYNPL--OGQKCIYOTTSMSQ 238
DB 179 --LGIDASEVELTRNNELIAVGKRSILKRLPFVFGMEPRILYNPLAHGOKCIVOTTSMSQ 236
QY 239 CSKTCGTGISTRTVNDNDECRVLETRICEVRPCGOPYSSILKKGKCSKTKKSPDEPRF 298
DB 237 CSKTCGTGISTRTVNDNDECRVLETRICEVRPCGOPYSSILKKGKCSKTKKSPDEPRF 296
QY 299 TYAGCLSVKTKRPRYKCGSCVDGRCTPOLTRTVKMRFCEDGERTFSKNVMMIIOCKCKNYN 358
DB 297 TYAGCLSVKTKRPRYKCGSCVDGRCTPOLTRTVKMRFCEDGERTFSKNVMMIIOCKCKNYN 356
QY 359 CPHANEAFPPYRLFNIDHKFRD 381
DB 357 CPHNEASFRLYSLFNIDHKFRD 379

RESULT 9
US-10-053-753-2
; Sequence 2, Application US/10053753
; Publication No. US20020150986A1
; GENERAL INFORMATION:

APPLICANT: Lau, Lester F.
TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/053.753
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28758/33766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: "Mouse Cyt61 amino acid sequence"
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-053-753-2

Query Match 91.2%; Score 1929; DB 14; Length 379;
Best Local Similarity 90.9%; Pred. No. 2.9e-157;
Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSSRIARALAVTLHLTRIALSTCPAACHCPLEAPKACAGVGLVRDGGCCCKVCAKOL 60
DB 1 MSSSTFRTLAVALVTLHLTRIALSTCPAACHCPLEAPKACAGVGLVRDGGCCCKVCAKOL 60
QY 61 NEDSKTOPCDHTGLEGCFNFGASSTALKGICRAOSEGRPCBYNSRITYONGESFOPNCKHQ 120
DB 61 NEDSKTOPCDHTGLEGCFNFGASSTALKGICRAOSEGRPCBYNSRITYONGESFOPNCKHQ 120
QY 121 CTCIDGAVGCIPLCPQELSLPNLGCPNRLVKTGQCCCEWVCDSDSIKIDMEDODGLG 180
DB 121 CTCIDGAVGCIPLCPQELSLPNLGCPNRLVKTGQCCCEWVCDSDSIKIDMEDODGLG 180
QY 181 KELGFDASEVELTRNNELIAVGKRSILKRLPFVFGMEPRILYNPL--OGQKCIYOTTSMSQ 238
DB 179 --LGIDASEVELTRNNELIAVGKRSILKRLPFVFGMEPRILYNPLAHGOKCIVOTTSMSQ 236
QY 239 CSKTCGTGISTRTVNDNDECRVLETRICEVRPCGOPYSSILKKGKCSKTKKSPDEPRF 298
DB 237 CSKTCGTGISTRTVNDNDECRVLETRICEVRPCGOPYSSILKKGKCSKTKKSPDEPRF 296
QY 299 TYAGCLSVKTKRPRYKCGSCVDGRCTPOLTRTVKMRFCEDGERTFSKNVMMIIOCKCKNYN 358
DB 297 TYAGCLSVKTKRPRYKCGSCVDGRCTPOLTRTVKMRFCEDGERTFSKNVMMIIOCKCKNYN 356
QY 359 CPHANEAFPPYRLFNIDHKFRD 381
DB 357 CPHNEASFRLYSLFNIDHKFRD 379

RESULT 10

US-09-853-625B-13
; Sequence 13, Application us/09853625B
; Patent No. US20020049304A1
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
; TITLE OF INVENTION: Human CCN-Like Growth Factor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
; CROCHT, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: us/09/853,625B
; FILING DATE: 14-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/053,587
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS: <Unknown>
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-853-625B-13
Query Match 80.1%; Score 1695; DB 9; Length 375;
Best Local Similarity 81.0%; Pred. No. 3,2e-137;
Matches 311; Conservative 20; Mismatches 41; Indels 12; Gaps 7;
QY 1 MSSRIAPALALVVTLLHLTRLAL-STCPAACHCPLAPKCAPGVGLVRDGGCGCKVCYAKQ 59
DB 1 MSAGAGAP-ALAAALICLARLALGSPCAVQCQPAAPQCAPGVGLVRDGGCGCKVCYAKQ 59
QY 60 LNEDESKTQPCDHTKLECFNFGASSTALKGICRAOSEGRCEYNSRIYONGSEFQNCQH 119
DB 60 LNEDESKTQPCDHTKLECFNFGASSTALKGICRAOSEGRCEYNSRIYONGSEFQNCQH 119
QY 120 QCTCIGAVGCIPLCPOELSLPMLGCPNPLVKTVOCCCEYVCDDESDIKDPMEDQDL 179
DB 120 QCTCIGAVGCIPLCPOELSLPMLGCPNPLVKTVOCCCEYVCDDESDIKDPMEDQDL 179
QY 180 GKELGFDASEVELTRNNELIAYGKGRSLKRLPVFGMEP--RLIYNPLQGGKCIYQTSMS 237
DB 178 SKREGLDASEGELTRNNELIAYKG-GIKMLPVFSGEPQSRAPENP----KCIYQTSMS 232
QY 238 QCSKTGTGISTVTNDNPECRIVKTRICEVRCQAPYYSILKKGKCKSKTKKSPSPYR 297
DB 233 QCSKTGTGISTVTNDNPECRIVKTRICEVRCQAPYYSILKKGKCKSKTKKSPSPYR 292
QY 298 FTYAGGLSVKYPKYGSCVDGRCTPOLTRTYKMRFRCEDEGTFSSKVMYMIOSCKY 357
DB 293 FTYAGGSSVKYKIRPKYKSCVDGRCTPOLTRTYKIRFRODDETFSTKSVMIOSCRNT 352
QY 358 NCPHANEAAFPFYRLFNDIHKFRD 381

DB 353 NCPHANE-AAFPFYRLVNDIHKFRD 375
RESULT 11
US-10-245-977-7
; Sequence 7, Application us/10245977
; Publication No. US20030113816A1
; GENERAL INFORMATION:
; APPLICANT: Weitz, Stephen L
; APPLICANT: Usinger, William R
; TITLE OF INVENTION: METHODS OF ASSAYING CONNECTIVE TISSUE GROWTH FACTOR
; FILE REFERENCE: EP0812 US
; CURRENT APPLICATION NUMBER: us/10/245,977
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 60/323,305
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-245-977-7
Query Match 45.1%; Score 955; DB 15; Length 347;
Best Local Similarity 46.2%; Pred. No. 6,9e-74;
Matches 176; Conservative 58; Mismatches 107; Indels 40; Gaps 7;
QY 1 MSSRIAPALALVVTLLHLTRLAL-STCPAACHCPLAPKCAPGVGLVRDGGCGCKVCYAK 58
DB 1 MASVAGPYSLALVLLCTRPATGDCSAQCCQCAABAARCPAGVSLVLDGGCGCKVCYAK 60
QY 59 QLNEDSKTQPCDHTKLECFNFGASSTALKGICRAOSEGRCEYNSRIYONGSEFQNCQH 118
DB 61 QLNEDSKTQPCDHTKLECFNFGASSTALKGICRAOSEGRCEYNSRIYONGSEFQNCQH 119
QY 119 HQCTCIDGAVGCIPLCPOELSLPMLGCPNPLVKTVOCCCEYVCDDESDIKDPMEDQDL 178
DB 120 YQCTCIDGAVGCIPLCPOELSLPMLGCPNPLVKTVOCCCEYVCDDESDIKDPMEDQDL 167
QY 179 LKELGFDASEVELTRNNELIAYGKGRSLKRL-PVFGMEPRLIYNPLQGGKCIYQTSMS 237
DB 168 -----KRTVVGPAALAVRLEDFEGPDPTM-----RANCIYQTSMS 205
QY 238 QCSKTGTGISTVTNDNPECRIVKTRICEVRCQAPYYSILKKGKCKSKTKKSPSPYR 297
DB 206 ACSKTGCMGISTVTNDNPECRIVKTRICEVRCQAPYYSILKKGKCKSKTKKSPSPYR 265
QY 298 FTYAGGLSVKYPKYGSCVDGRCTPOLTRTYKMRFRCEDEGTFSSKVMYMIOSCKY 357
DB 266 FELSGCTSVKTYAKKCGVCTDGRCTPRTTLLPVFPCDSEIKNKMMFIKTCACHY 325
QY 358 NCPHANEAAFPFY--RLFNDI 376
DB 326 NCPGDNDIFESLYRRMYGDM 346
RESULT 12
US-10-101-040-3
; Sequence 3, Application us/10101040
; Publication No. US20020142353A1
; GENERAL INFORMATION:
; APPLICANT: FIBROGEN, INC
; APPLICANT: SCHMIDT, Brian
; APPLICANT: ALLEN, Margaret
; APPLICANT: SVERDRUP, Fran
; APPLICANT: CARMICHAEL, David
; TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR (CTGF) AND METHODS OF
; FILE REFERENCE: FIBRO100-1
; CURRENT APPLICATION NUMBER: us/10/101,040
; CURRENT FILING DATE: 2002-03-18

PRIOR APPLICATION NUMBER: 09/292,036
PRIOR FILING DATE: 1999-04-14
PRIOR APPLICATION NUMBER: US 09/292,036
PRIOR FILING DATE: 1999-04-14
PRIOR APPLICATION NUMBER: US 09/187,478
PRIOR FILING DATE: 1998-11-06
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 348
TYPE: PRT
ORGANISM: Mouse
US-10-101-040-3

Query Match 45.0%; Score 952.5; DB 14; Length 348;
Best Local Similarity 46.6%; Pred. No. 1.1e-73;
Matches 178; Conservative 57; Mismatches 106; Indels 41; Gaps 8;

QY 1 MSRIARALAVTLVTLHL-TRLAL-STCPACHCPLF-APKCAPGVGLVRDGGCCGYCA 57
DB 1 MIAVAGPISIALVIALCTRPATGDCSACQCAAEAPHCPRAGVSLVLDGCGCCRYCA 60
QY 58 KQLNEDSKTOPCDHTKLECNFGASSTALKGICRAQSEGRPCVNSRIYONGESFOPNC 117
DB 61 KQLGELCTERDPCDHPKGLFCDFGSPANKRIGVCTAK-DGAPCVGGSVYRSGESFQSSC 119
QY 118 OHQCTCIDGAVGCIPLCQELSLPRLGCPNRLVYVQCCCEVWCDEDSIKDPEDDGG 177
DB 120 KYQCTCLDGAAGVCLSMVRLPSDPCPPRRVRLPKQCKCEWVCDP----- 168
QY 178 LIGKELGPDASEVELTRNNELIANGKGRSLKRL-PVFGMEPRILYNPLOGCKIYQTSW 236
DB 169 -----KQRTAVGPALAAARLEDTGPDPTM-----RANCLVQTTW 205
QY 237 SQSKTCGTGISTRTVNDNPECRIVKETRICEVRCGQPVYSSLKGRKSKTKSPBPV 296
DB 206 SACSCTGCMGISTRTVNDNFCRLKQSLCMVRPCADLEENIKKGRKCIPTPKIAPV 265
QY 297 RFTYAGCLSVKRYKRYCGSVDCRCCTPOLTRVYKMFREDGETFSKNVMMIOQSKCN 356
DB 266 KFLSGCTSVKRYAKRCGVCTDRCCTPHRTTLLPVEFKCPDGEIMKNNMFKTQACH 325
QY 357 YNCPHANEAFPEY--RLFNDI 376
DB 326 YNCPGNDIFESLRYRKNYGD 347

RESULT 13
US-09-853-625B-15
Sequence 15, Application US/09853625B
Patent No. US20020049304A1
GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
CROCHT, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/853,625B
FILING DATE: 14-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/053,587
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: MOLLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS: <Unknown>
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-853-625B-15

Query Match 44.8%; Score 948.5; DB 9; Length 348;
Best Local Similarity 46.3%; Pred. No. 2.5e-73;
Matches 177; Conservative 58; Mismatches 106; Indels 41; Gaps 8;

QY 1 MSRIARALAVTLVTLHL-TRLAL-STCPACHCPLF-APKCAPGVGLVRDGGCCGYCA 57
DB 1 MIAVAGPISIALVIALCTRPATGDCSACQCAAEAPHCPRAGVSLVLDGCGCCRYCA 60
QY 58 KQLNEDSKTOPCDHTKLECNFGASSTALKGICRAQSEGRPCVNSRIYONGESFOPNC 117
DB 61 KQLGELCTERDPCDHPKGLFCDFGSPANKRIGVCTAK-DGAPCVGGSVYRSGESFQSSC 119
QY 118 OHQCTCIDGAVGCIPLCQELSLPRLGCPNRLVYVQCCCEVWCDEDSIKDPEDDGG 177
DB 120 KYQCTCLDGAAGVCLSMVRLPSDPCPPRRVRLPKQCKCEWVCDP----- 168
QY 178 LIGKELGPDASEVELTRNNELIANGKGRSLKRL-PVFGMEPRILYNPLOGCKIYQTSW 236
DB 169 -----KQRTAVGPALAAARLEDTGPDPTM-----RANCLVQTTW 205
QY 237 SQSKTCGTGISTRTVNDNPECRIVKETRICEVRCGQPVYSSLKGRKSKTKSPBPV 296
DB 206 SACSCTGCMGISTRTVNDNFCRLKQSLCMVRPCADLEENIKKGRKCIPTPKIAPV 265
QY 297 RFTYAGCLSVKRYKRYCGSVDCRCCTPOLTRVYKMFREDGETFSKNVMMIOQSKCN 356
DB 266 KFLSGCTSVKRYAKRCGVCTDRCCTPHRTTLLPVEFKCPDGEIMKNNMFKTQACH 325
QY 357 YNCPHANEAFPEY--RLFNDI 376
DB 326 YNCPGNDIFESLRYRKNYGD 347

RESULT 14
US-10-053-753-6
Sequence 6, Application US/10053753
Publication No. US20020150986A1
GENERAL INFORMATION:
APPLICANT: Iau, Lester F.
TITLE OF INVENTION: Extracellular Matrix Signalling Molecules
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Marshall, O'Toole, Garsteln, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/10/053,753
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Coughn, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 28758/33766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: "P1sp12 amino acid sequence"
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-053-753-6

```


Dp 179 -- LGEDASEVELTRNNELIAIGKSSIKRLPYFGTEPRVLNPLHAHGQKCIQYTTSMWQ 23

submitted to the EMBL Data Library, October 1998

A:Reference number: Z20293

A:Accession: T26972

A:Status: preliminary; translated from GB/EMBL/DDat

A:Molecule type: DNA

A:Residues: 1-1111 <MIL>

A:Cross-references: EMBL:AL033657; PIDN:CAA21739.1; GSPDB:GNO0019; CESP:Y47H9C.4

A:Experimental source: clone Y47H9C

C:Genetics:

A:Gene: CESP:Y47H9C.4

A:Map position: 1

A:Insertions: 50/2; 84/2; 150/1; 238/3; 342/3; 797/1; 851/1; 947/2; 1017/1; 1083/1

C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology

Query Match

8.2%; Score 174; DB 2; Length 1111;

Best local similarity 19.7%; Pred. No. 4.3e-05;

Matches 84; Conservative 33; Mismatches 161; Indels 148; Gaps 17;

```

QY 26 CPAACHCPLE-APKAPGVGLVR-----DGGGCKVCAR-QLNEDCSKTQPCDHTKGLGEC 78
DB 380 CSKTCCTCCEBENTLMCAPNTGFCRCRKGFGDNCCLASCKDSYGPNCCKQAMCDMNHASEC 439
QY 79 NGASSTALKGICRAQSEGRPCENSRITONGESRQPNQHOCTC-----IDGA----- 127
DB 440 NPETGSC---VCKPRTGKNCSEPCPL---DFGPNCAHQCCQCMORGVCYCGADGCKCO 491
QY 128 -----VGCIPPL-----CPQELSLPN--LGCP-- 146
DB 492 CDKGTGTHRCHEHHPADTFGANCEKRCRCKGICGCPITGECGCPAGLOGANDIDGCPGEG 551
QY 147 -----NPLRVKVTGQCCCEWVCDEDSIKDPMEDDGLGKELGPDASEVELT 193
DB 552 SYPGCKLHKCKVCVNGCDKDETEGC---TC-----QPGFGSCDSTTCSKGYG 596
QY 194 RNNELIYAVGKGRSLKRLPFQGMERILYNPLQGQKCIQVOTSSQCSKTCGTISRVTN 253
DB 597 ESECL-----SCPCSDASCSKOTGKCLPLGTRGVS 627
QY 254 DNEPCRLVETRICVEYRPGQPYVSSIKKCKSKTRKSPPEVFTY-----AGCLS 305
DB 628 CDQKCDPNTFGLQGTVPSPCASIDPNKGVCLSPGSSGJHCEHNCVAGSYGQCCQ 667
QY 306 VKYRKYKGCSCYDGCCTPOLTRIV-----KRFRCDEGETSKNVMITQSCKNY 357
DB 688 V-----C-SCADHGCDPPTGECICDEPGYHGKTCSEKCPDKYGVGALDCKPCASGS 739
QY 358 NCPHAN 363
DB 740 TCDHIN 745

```

RESULT 8

T27283

hypothetical protein Y64G10A.f - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27283

R:Almscough, R.

submitted to the EMBL Data Library, September 1999

A:Reference number: Z20336

A:Accession: T27283

A:Status: preliminary; translated from GB/EMBL/DDat

A:Molecule type: DNA

A:Residues: 1-1620 <MIL>

A:Cross-references: EMBL:AL110498; NID:EL542303; PIDN:CA854471.1; CESP:Y64G10A.f

A:Experimental source: clone Y64G10A

C:Genetics:

A:Gene: CESP:Y64G10A.f

A:Insertions: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 548/1; 559/1; 601/1; 625/1;

Query Match

8.2%; Score 173.5; DB 2; Length 1620;

Best local similarity 24.2%; Pred. No. 6.8e-05;

Matches 92; Conservative 33; Mismatches 146; Indels 109; Gaps 21;

```

QY 26 CPAACHCPLEAP-----KCAPG-----VGLVRDGGCCCKVCARQALNEDCSK 66
DB 1052 CKGICCCQNGATDSDVTSVSGECRPGWRGKKRCDRCPDG--RFEGCNALICDCTTNDTSM 1109
QY 67 TOP----CDHTKLECNFGASSTALKGICRAQSEGRPCENSRITONGESRQPNQHOCT 122
DB 1110 YNPFVARCHVHTG-EGR-----CPAGTGDCCQTSCLPLHGG--GCHHSQ 1153
QY 123 CIDGAVGCIPLCPQELSLPNLGGPNRLVKTGQCCCEWVCDEDSIKDPMEDDGLGKE 182
DB 1154 CSNGA-----SCRYVTGF--CDPSPGPMNCSECEPGLMGSCMKHCLMHGECNKE 1206
QY 183 LGFDASEVELTRNNELIYAVGKGRSLKRLPFQGMERILYNPLQGQKCIQVOT-----TSW 236
DB 1207 NG-----DCEIDGWTGSLDPFGQFGNCAQRCKNKGASCDRRTGCECLPGW 1256
QY 237 S--QCSKTGTGTISTVTNDNPECLVLETRICEVYRPGQPYVSSLK-----KGRKCSKT 289
DB 1257 SGHECHERKSCVSG-----HYGANC---EFTCEBENCALCDPLISGHSCQPGWGRKRCNR 1307
QY 290 KSPPEVRYFTYAGCLSVKRYRKYCG-----SC--VDGRCTPOLTRVYKMRFRCE 338
DB 1308 -----CL--KGYFGRHCSQSCRCANSKCDHISGRQCPKGYAGHSTELCP 1352
QY 339 DGETFSKNVMITQSCKNYN 358
DB 1353 DG-TFEGSCS--QKDCGGEN 1369

```

RESULT 9

S08167

Baldiant ring 3 protein - midge (Chironomus tentans)

C:Species: Chironomus tentans

C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Jul-2000

C:Accession: S08167

R:Paulsson, G.; Lendahl, U.; Galli, J.; Ericsson, C.; Wieslander, L.

J. Mol. Biol. 211, 331-349, 1990

A:Title: The baldiant ring 3 gene in Chironomus tentans has a diverged repetitive str

A:Reference number: S08167; MID:9012404; PMID:1689777

C:Accession: S08167

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-1700 <PAU>

A:Cross-references: GB:X52263; NID:97057; PIDN:CAA36506.1; PID:97058

C:Genetics:

A:Gene: BR3

A:Map position: 4

C:Superfamily: unassigned Baldiant ring proteins

Query Match

7.4%; Score 157; DB 2; Length 1700;

Best local similarity 20.9%; Pred. No. 0.0012;

Matches 88; Conservative 47; Mismatches 170; Indels 116; Gaps 21;

```

QY 30 CHCPLEAPK--CAPGYGLVRDGGC--CKYCAKOLNEDCSKTQPCDHTKGLGEC-----NFG 81
DB 221 CICTPTEPAGGCSAPLKWDDKSCACAPKMEKKEKCVESGRTWNPNTCEGCAQLNCP 280
QY 82 ASSTALKGICRAQ--SEGRPCENSRITONGESRQPNQHOCTCI----- 124
DB 281 DNKKANKKTCQCCCKEYKCC-----NGQVF--CKDSGCVCGGDKDKTCTAPQY 330
QY 125 DGAVGCIPLCPQELSLPNLGGPNRLVKTGQCCCEWVCDEDSIKDPMEDDGLGKEIG 184
DB 331 DG-VACSGCPVYVMQKRPADGCPRP-----QKWDRECHCECPVK-HDCKNKGKVM 378
QY 185 FDASEVELTRNNELIYAVGGR-----SLKRLPFQGMERILYNPLQGQKCIQVOTSM 236
DB 379 ETIQGQICPRDAPVCTAGKERGCESECKCINNEPREGAKPLWNE-NTCKCVCPADK- 436
QY 237 SQGSKTCGTGISRVTNDNDEC-----RLVKEETRICVVR-----GQAPYSSSL 280
DB 437 QMSPGGCGSGKSNKLTTCQCECDQASAKGLRWMADTKCKECQPGMPGEGCGKQWISD 496

```

```

OY 281 KKGGKSKT-----KSPSPVPTVAGLCKSVKKYRP--- 311
      ||| |
Db 497 KCKECSPTITCOAPQIILINTECKCPVNLAKCKSPKOWTDSKCLCESTTPATC 556
      ||| |
OY 312 ----KYCG---SCV---DGRCTPOLTRTKMFRCEDEET-FSKVMMIIOGCKNYNC 359
      ||| |
Db 557 EKGOTWGEACQICIGPDGDKNGKMKFPDPSCECKCKNNPTCTSPVMMADDECC--KC 614
      ||| |
OY 360 P 360
      ||| |
Db 615 P 615

RESULT 10
T13954
MEGF6 protein - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T13954
R:Nakayama, M., Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
Genomics 31, 27-34, 1998
A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
A:Reference number: Z14126; MUID:98360089; PMID:9693030
A:Accession: T13954
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1574 <NA>
A:Cross-references: EMBL:AB011532; NID:g3449293; PIDN:BAA32462.1; PID:g3449294
A:Experimental source: strain Sprague-Dawley; brain
C:Genetics:
A:Gene: MEGF6

Query Match 7.28; Score 153; DB 2; Length 1574;
Best Local Similarity 21.94; Pred. No. 0.0021;
Matches 97; Conservative 36; Mismatches 133; Indels 176; Gaps 28;

OY 26 CPACHCP-----LEAPKCAPGV--GLVRDGC----- 50
      ||| |
Db 568 CSSECTQONGCTDPVLAGACRCPGVSGAHCEGCEPKGYGKCKKCHCANNRCHRLY 627
      ||| |
OY 51 GCCKVCAPKOL-----NEDCKTOPCDHTKGLCNFGASSTALGICR 92
      ||| |
Db 628 GAC-ICDPGLYGRFCHLACPPMAFGPCSDCLCEQ--SHTRSCNPRDSCS-----CK 678
      ||| |
OY 93 AQBEGRPCEYNSTRIVNGESFOPNCOHCTCIDGAVGCIPL-----CPQ 136
      ||| |
Db 679 AGFQGERCAEC---ESG--FEGPGCRHRTCTCPG--VACDPVSGECRTQCPGYGEGDCGQ 733
      ||| |
OY 137 ELSLPNG-----CPNRLVKNYGC--CEMWCDDESDIKPMEDQDILLKELGFDA 187
      ||| |
Db 734 ECPVGTGTVNGSSGSCVGPACHRVTEGCLCPKGTGEDGACDPESRWGLGCOEI--- 789
      ||| |
OY 188 SEVELFRNNELIIVAGKRSILKRLPVFQMEPRILYNPLQOGKCIYQTSMSQSKTC--- 243
      ||| |
Db 790 -----CPACRCHASC-----NPTGTGCLCLPGVVSRCQDITCSAGW 825
      ||| |
OY 244 -GTGISTR--VTNDNPECLRVKETRICEVRPCQPVYSILKKGKCKSKTKSPPEVFTY 300
      ||| |
Db 826 YGTGCOIRACACANDG-HCDPTTGRSC-----APGWTGLSCORACDSGHMGDP----- 872
      ||| |
OY 301 AGTLVKKYKRYC---GSC--VDGRC-CT-----POLRLTYKMRP---RCDEGEFFSK 345
      ||| |
Db 873 --CI-----HPCMSAGHGKDAVSGCLCEAGIEGRCHOSCKGQYGPSC----- 918
      ||| |
OY 346 NVMMIIOGCKNYN--CPHANEA 365
      ||| |
Db 919 -----QKCRCEHGAACDHVSGA 935
      ||| |

RESULT 11
T42215
zonadhesin - mouse

```

```

N:Alternate names: sperm-specific membrane protein
C:Species: Mus musculus (house mouse)
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T42215
R:Gao, Z.; Garbers, D.L.
J. Biol. Chem. 273, 3415-3421, 1998
A:Title: Species diversity in the structure of zonadhesin, a sperm-specific membrane
A:Reference number: Z22080; MUID:98123114; PMID:9452463
A:Accession: T42215
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-5376 <GAO>
A:Cross-references: EMBL:U07068; NID:g3327420; PID:g3327421; PIDN:AAC26680.1
C:Genetics:
A:Gene: Zan
A:Map position: 5
C:Function:
A:Description: functions in multiple cell adhesion processes
A:Note: found exclusively on the apical region of the sperm head
C:Keywords: cell adhesion

Query Match 7.28; Score 152; DB 2; Length 5376;
Best Local Similarity 21.08; Pred. No. 0.0082;
Matches 81; Conservative 39; Mismatches 138; Indels 128; Gaps 19;

OY 19 TRIALSTCPACHCPLEAPKCAPGVGLVRDGCCKVCANQINEDCKTOPCDHTKGLGEC 78
      ||| |
Db 4378 TKIALQ-CPAHSHTTYCLPSCITFSCSNVNDRC-----ESTLQRTSTICGLC 4425
      ||| |
OY 79 NFGASSTALGICRAQ-----SEGRPEYNSRIYONGESF-QPNCOHCTCIDGAVGCIPL 132
      ||| |
Db 4426 HSGFVFEKDKCVPRTOGGCKDSQ-----TLIPAKNMWTTGCSQRCCTGGLVQCHD 4478
      ||| |
OY 133 L-CPQELSLPMLGCPNRLVKNYGC-----CE-----EMVCDDESDIKPME 173
      ||| |
Db 4479 FOCPSGAECDDIEDGNSNCVEITVQCPAHSYSKCLPPCQPSCDPDGHEGTSPEAPST 4538
      ||| |
OY 174 DQDGLKELGELFQDSEVELFRNNELIIVAGKRSILKRLPVFQMEPRILYNPLQOGKCIYQT 233
      ||| |
Db 4539 CEECV-----CEPDYV--LSNDKCVPS- 4559
      ||| |
OY 234 TSMQSKCTGCTGISTRVTNDNPECLRVKETRICEVRPC-----GQPVYS--SLKKG 283
      ||| |
Db 4560 ---SECCCKNAHG-----LIPESKTVMSGCKTKCKCKGTVQCHDFSCPTG 4604
      ||| |
OY 284 KCKSKTKSPB-----PVRETYAGCLSVKKYKRYGSCVD--GRC--CTPOLTRT 330
      ||| |
Db 4605 SRCIDNNEGNSNCVTYALKCPAHSILYTNCL-----PCLPSCSDPDGLCGTSPPEVPT 4658
      ||| |
OY 331 VKMFRCEDETFESKNYMTQ--SCCK 355
      ||| |
Db 4659 CKEGCIQSGVYLHKRCKMLRHCDG 4684
      ||| |

RESULT 12
A39804
thrombospondin precursor - chicken
C:Species: Gallus gallus (chicken)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A39804
R:Lawler, J.; Dugnette, M.; Ferro, P.
J. Biol. Chem. 266, 8039-8043, 1991
A:Title: Cloning and sequencing of chicken thrombospondin.
A:Reference number: A39804; MUID:91217026; PMID:2022631
A:Accession: A39804
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1178 <LAW>
A:Cross-references: GB:M60853; NID:g212763; PIDN:AAA51437.1; PID:g212764
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology;
F:325-383/Domain: von Willebrand factor type C repeat homology <VWC>
F:386-437/Domain: thrombospondin type 1 repeat homology <THRI>
F:442-498/Domain: thrombospondin type 1 repeat homology <THR2>

```


notch4 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 08-Sep-2002

C:Accession: T09059

R:Rowen, L.; Mahairas, G.; Qin, S.; Ahearn, M.E.; Dankers, C.; Lasky, S.; Loretz, C.; Sc submitted to the EMBL Data Library, October 1997

A:Description: Sequence of the mouse major histocompatibility locus class III region.

A:Reference number: 216543

A:Accession: T09059

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1964 <ROW>

A:Cross-references: EMBL:AF030001; NID:92564945; PID:92564947

C:Genetics:

A:Gene: notch4

A:Map position: 17

A:Introns: 22/1; 49/2; 148/1; 264/1; 305/1; 384/1; 436/1; 501/1; 539/1; 577/1; 618/1; 67

1679/3; 1729/1; 1761/3

C:Superfamily: notch protein; ankyrin repeat homology; EGF homology

C:Keywords: receptor; signal transduction

F:514-545/Domain: EGF homology <EGF>

Query Match 6.9%; Score 146; DB 2; Length 1964;

Best Local Similarity 23.4%; Pred. No. 0.0086;

Matches 94; Conservative 28; Mismatches 141; Indels 138; Gaps 24;

```
QY 28 AACGCPLEAPK--CAPGVGLVWDGCGCCCKVCAKOLNEDCKT-----OP-----CD 71
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 523 AACHDILNGFQCLDPFTGAR-----CEKDME--CSSTPCANGRCRDQPAFYCE 573
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 72 HTKGLA-----CNGASTALKG---ICRAQSEGRPEY---NSRIYON 109
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 574 CLPGEHPHCEKEVDECLSDPCPVGASCLDLPGAFFCLCRPGFTGOLCEVPLCTPNMOP 633
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 110 GESFPNCOHQ-----CTCIDGAVGCIPL-----CPQELSLPNLGCNPRLVAVT 154
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 634 GQ-----QCQGEHRAAPCLCPDGSFGCPAEDMCPCHHGCQSLCVCDEGWTGPBCETEL 689
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 155 GOCCEEWVCDDEDSIKDPMEDODGLGKELGFDASEVELTRNNELIYAVGKRLKRLPVFG 214
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 690 GGCIST-PCAHGTCCHPOPS-----GYNCTCPAGYMGILCSEVTACHSGPC----- 735
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 215 MEPRILYNPLQOKCTIVOTTSMSQCSKTGCTGISTVTNDNPECLVKEETRIE--VRPC 272
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 736 -----LNGGSCSIREGYSCITLPSHTG-----RHQQTAVDHC 768
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 273 GQPVYSLKKKKCKSKTKSPKSPVFTYAGCLSVKKYRPKYC-----GSCVDGRC----- 322
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 769 ---VSASCLNGTCVN-----KQITF---CLCATGFQGLHCEKTNPSCADSPCRNKAT 817
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 323 C--TPQLTRYVKMRFCEDETFSSKNVMTQSCCKNTNCPH 361
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 818 QDPTPRGARCL-----CSPGTGSSCQTLIDLQ-ARRPCPH 852
   |||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
```

Search completed: August 5, 2003, 14:07:29

Job time : 17.5408 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:02:16 ; Search time 9.525 Seconds

(without alignments)
1881.066 Million cell updates/sec

Title: US-09-495-448A-4
Perfect score: 2116
Sequence: 1 MSSRIARALALVLTLLHTR.....ANEAPPFRLNDHKFND 381

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2106	99.5	381	1 CYR6_HUMAN	P00622 homo sapien
2	1929	91.2	379	1 CYR6_MOUSE	P18406 mus musculu
3	1929	91.2	379	1 CYR6_RAT	O96872 rattus norv
4	1695	80.1	375	1 CE10_CHICK	P19336 gallus gall
5	955	45.1	347	1 CTGF_RAT	O91899 rattus norv
6	952.5	45.0	346	1 CTGF_MOUSE	P29268 mus musculu
7	947	44.8	349	1 CTGF_HUMAN	P29279 homo sapien
8	933	44.1	349	1 CTGF_BOVIN	O18739 bos taurus
9	906	42.8	349	1 CTGF_PIG	O19113 sus scrofa
10	851.5	40.2	351	1 NOV_CHICK	P28686 gallus gall
11	848	40.1	353	1 NOV_CHICK	P42642 columba co
12	827.5	39.1	357	1 NOV_HUMAN	P48745 homo sapien
13	825.5	39.0	343	1 NOV_XENIA	P51609 xenopus lae
14	822	38.8	354	1 NOV_MOUSE	O64299 mus musculu
15	818.5	38.7	351	1 NOV_RAT	O96295 rattus norv
16	523.5	24.7	250	1 CTGL_HUMAN	O76076 homo sapien
17	522.5	24.7	251	1 CTGL_MOUSE	O92094 mus musculu
18	490	23.2	250	1 CTGL_RAT	O91866 rattus norv
19	157	7.4	1700	1 BAR3_CHITE	O03376 chironomus
20	153.5	7.3	2282	1 ZAN_RABIT	P57999 oryctolagus
21	152	7.2	3110	1 LMA2_HUMAN	P24043 homo sapien
22	152	7.2	5376	1 ZAN_MOUSE	O68789 mus musculu
23	150.5	7.1	1178	1 TSP2_CHICK	P35440 gallus gall
24	150	7.1	837	1 MUC4_RAT	P98089 rattus norv
25	147.5	7.0	456	1 NRTL_CHICK	O90265 gallus gall
26	146	6.9	1964	1 NTC4_MOUSE	P31695 mus musculu
27	144.5	6.8	1172	1 LMB3_HUMAN	O13751 homo sapien
28	143	6.8	1173	1 TSP1_XENIA	P35448 xenopus lae
29	141	6.7	480	1 HRA1_MOUSE	O9118 mus musculu
30	141	6.7	1233	1 MUC2_HUMAN	P98088 homo sapien
31	141	6.7	5179	1 MUC2_HUMAN	O92081 mus musculu
32	139	6.6	447	1 NRTL_MOUSE	O920C1 mus musculu
33	139	6.6	1170	1 TSP1_BOVIN	O28178 bos taurus

34	139	6.6	1587	1 LMG3_HUMAN	O9566 homo sapien
35	139	6.6	2437	1 NTC1_BRARE	P46530 brachydanio
36	139	6.6	2482	1 VWF_PIG	O28833 sus scrofa
37	138.5	6.5	2703	1 NOTC2_DROME	P07207 drosophila
38	137.5	6.5	1504	1 SLIT1_DROME	P24014 drosophila
39	137	6.5	453	1 HRA3_HUMAN	P83110 homo sapien
40	137	6.5	870	1 SRC2_HUMAN	O96966 homo sapien
41	137	6.5	2531	1 NTC1_RAT	O07008 rattus norv
42	136.5	6.5	833	1 SRC2_MOUSE	P59222 mus musculu
43	135	6.4	1168	1 LMB3_MOUSE	O61087 mus musculu
44	135	6.4	1170	1 TSP1_MOUSE	P35441 mus musculu
45	135	6.4	2813	1 VWF_CANFA	O28295 canis fami

ALIGNMENTS

RESULT 1
CYR6_HUMAN STANDARD: PRT: 381 AA.
AC 000622; 014934; 043775; 09BZL7;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)
DE (Insulin-like growth factor-binding protein 10) (IGF1 protein).
GN CYR61 OR IGFBP10 OR GIG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Albrecht C., von der Kammer H., Klaudiny J., Mayhaus M., Nitsch R.M.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97280750; PubMed=9135077;
RA Juy P., Berge-Leffranc J.L., Marsollier C., Mejean C., Taviaux S.,
RA Berta P.;
RT "The human growth factor-inducible immediate early gene, CYR61, maps
RT to chromosome 1p.";
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98197344; PubMed=9536281;
RA Martinierie C., Viegas-Pequignot E., Nguyen V.C., Perbal B.;
RT "Chromosomal mapping and expression of the human cyr61 gene in tumour
RT cells from the nervous system.";
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Kolesnikova T.V., Lau L.F.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Bl A.B., Yu L.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Schuetz N., Lechner A., Groll C., Koehrl J., Jakob F.;
RT "Regulation of hCYR61 by vitamin D, serum and cytokines in fetal human
RT osteoblasts.";
RN [7]
RP SEQUENCE FROM N.A.
RA Leng E., Tai G., Estable M., Liu J., Chow C., Sadowski I.;
RT "Organization and expression of the CYR61 gene in normal human
RT fibroblasts.";
RN [8]
RP Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.

RC TISSUE-Lung, Placenta, and Skin;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carrincci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Halys S.W.,
 RA Villalón D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield T.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
 CC PROTEINS. PROMOTES PROLIFERATION, MIGRATION AND ADHESION (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 WFCC domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Y12084; CAA72802.1; -;
 DR EMBL; U62015; AAB58319.1; -;
 DR EMBL; Y11307; CAA72167.1; -;
 DR EMBL; AF003594; AAB61240.1; -;
 DR EMBL; AF031385; AAB84227.1; -;
 DR EMBL; Z98053; CAB10848.1; -;
 DR EMBL; AF307860; AAG59863.1; -;
 DR EMBL; BC001271; AAH01271.1; -;
 DR EMBL; BC009199; AAH09199.1; -;
 DR EMBL; BC016952; AAH16952.1; -;
 DR GeneW; HGNC:2634; CYR61.
 DR MIM; 602369; -;
 DR GO; GO:0008283; P:cell proliferation; TAS.
 DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Insu_gro_fac-pr.
 DR InterPro: IPR001007; WVF_C.
 DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00019; IGFBP; 1.
 DR Pfam; PF00090; TSP_1; 1.
 DR Pfam; PF00093; WVF; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; WVC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF_BINDING; 1.
 DR PROSITE; PS00992; TSP1; 1.
 DR PROSITE; PS01208; WVF_C; 1.
 DR PROSITE; PS0184; WVF_C_2; 1.
 DR Growth factor binding; Signal.

FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 381 CYR61 PROTEIN.
 FT DOMAIN 26 97 IGFBP.
 FT DOMAIN 98 164 WFCC.
 FT DOMAIN 228 273 TSP TYPE-1.
 FT DOMAIN 286 360 CTCK.
 FT DISULFID 286 323 BY SIMILARITY.
 FT DISULFID 303 337 BY SIMILARITY.
 FT DISULFID 314 353 BY SIMILARITY.
 FT DISULFID 317 355 BY SIMILARITY.
 FT DISULFID 322 359- BY SIMILARITY.
 FT CONFLICT 165 165 E -> Q (IN REF. 3).
 FT CONFLICT 210 210 L -> I (IN REF. 5).
 FT CONFLICT 220 220 L -> R (IN REF. 5).
 FT CONFLICT 369 369 F -> L (IN REF. 7).
 SO SEQUENCE 381 AA: 42026 MW: 42026 MW: FC0BD39C078CA0B1 CRC64;
 Query Match 99.5%; Score 2106; DB 1; Length 381;
 Best Local Similarity 99.5%; Pred. No. 9.8e-146;
 Matches 379; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSSRIARALAVTLHLTRIALSTCPACHCPLPAKCAPGVGLVRDGGCGCKYCAKQL 60
 DB 1 MSSRIARALAVTLHLTRIALSTCPACHCPLPAKCAPGVGLVRDGGCGCKYCAKQL 60
 QY 61 NEDSKTOPCDHTKGLKCNFGASSTALKGICRAOSEGRPCERYNSRTYONGSEFOPNCQH 120
 DB 61 NEDSKTOPCDHTKGLKCNFGASSTALKGICRAOSEGRPCERYNSRTYONGSEFOPNCQH 120
 QY 121 CTCIDGAVGCIPLCPQOELSLNLCGPNRLKLVGGCCCEWVCDSDSKIDMEDODGLG 180
 DB 121 CTCIDGAVGCIPLCPQOELSLNLCGPNRLKLVGGCCCEWVCDSDSKIDMEDODGLG 180
 QY 181 KELGDAEVELTRNNELIANGKSKRLPVGNEPRILYNPLGOKCIVOTTSWQCS 240
 DB 181 KELGDAEVELTRNNELIANGKSKRLPVGNEPRILYNPLGOKCIVOTTSWQCS 240
 QY 181 KELGDAEVELTRNNELIANGKSKRLPVGNEPRILYNPLGOKCIVOTTSWQCS 240
 DB 241 KTCGIGSTRYTNNDPEKRLKERRICVPRCGQPVYSSLKKGKCKSTKSPPEVRYTY 300
 QY 241 KTCGIGSTRYTNNDPEKRLKERRICVPRCGQPVYSSLKKGKCKSTKSPPEVRYTY 300
 DB 241 KTCGIGSTRYTNNDPEKRLKERRICVPRCGQPVYSSLKKGKCKSTKSPPEVRYTY 300
 QY 301 AGCLSVKKYRPKYGSCVGDGRCCTPOLRTYKMRRCDEGFENVMYIOSCKCNVNC 360
 DB 301 AGCLSVKKYRPKYGSCVGDGRCCTPOLRTYKMRRCDEGFENVMYIOSCKCNVNC 360
 QY 361 HANEAPFPYRLFNDIHKFRD 381
 DB 361 HANEAPFPYRLFNDIHKFRD 381
 RESULT 2
 CYR6_MOUSE STANDARD; PRT; 379 AA.
 ID CYR6_MOUSE
 AC P18406;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)
 GN (Insulin-like growth factor-binding protein 10) (3CH61).
 GN CYR61 OR IGFBP10.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP STRAIN=BA1B/C; TISSUE=Fibroblast;
 RC MEDLINE=90287146; PubMed=2355916;
 RX O'Brien T.P., Yang G.P., Sanders L., Lau L.F.;
 RT "Expression of cyr61, a growth factor-inducible immediate-early
 RT gene.";
 RL Mol. Cell. Biol. 10:3569-3577(1990).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN-AJ; TISSUE=Embryonic fibroblast;
 RX MEDLINE=91288203; PubMed=2062642;
 RA Latickie B.V., O'Brien T.P., Lau L.F.;
 RT "Promoter function and structure of the growth factor-inducible
 RT immediate gene cyr61."
 RL Nucleic Acids Res. 19:3261-3267(1991).
 CC -1- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
 CC PROTEINS; PROMOTES PROLIFERATION, MIGRATION AND ADHESION.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: LOW IN KIDNEY, ADRENAL GLAND, TESTES, BRAIN,
 CC AND OVARY, MODERATE IN HEART, UTERUS, AND SKELETAL MUSCLE, HIGHEST
 CC IN LUNG.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED FROM G(0)/G(1) THROUGH MID-G(1) IN
 CC NORMAL CELLS. AND AT A CONSTANT LEVEL IN RAPIDLY GROWING CELLS.
 CC -1- INDUCTION: By growth factors.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 WFPC domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M32490; AAA37512.1; -;
 DR EMBL; X56790; CAA40109.1; -;
 DR PIR; A35669; A35669;
 DR MGI; MGI:88613; Cyr61.
 DR GO; GO:0001569; Patterning of blood vessels; IMP.
 DR InterPro; IPR006208; Cys_knot.
 DR InterPro; IPR006207; Ins1_knot.C.
 DR InterPro; IPR000867; Ins1_gro_fac-pr.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR001007; WFC_C.
 DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00219; IGFBP; 1.
 DR Pfam; PF00090; TSP_1; 1.
 DR Pfam; PF00093; WFC; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; WFC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00223; IGF_BINDING; 1.
 DR PROSITE; PS50093; TSP1; 1.
 DR PROSITE; PS01208; WFC_1; 1.
 DR PROSITE; PS50184; WFC_2; 1.
 DR PROSITE; PS50184; WFC_2; 1.
 KW Growth factor binding; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 379 CYR61 PROTEIN.
 FT DOMAIN 98 164 WFPC.
 FT DOMAIN 226 271 TSP TYPE-1.
 FT DOMAIN 284 358 CTCK.
 FT DISULFID 301 321 BY SIMILARITY.
 FT DISULFID 312 351 BY SIMILARITY.
 FT DISULFID 315 353 BY SIMILARITY.
 FT DISULFID 320 357 BY SIMILARITY.
 SQ SEQUENCE 379 AA; 41709 MM; FA6B5014B56A8EE9 CRC64;

Query Match 91.2%; Score 1929; DB 1; Length 379;
 Best Local Similarity 90.9%; Pred. No. 6.5e-133;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

OY 1 MSRIARALAVVTLHLTRIALSTCPACHCPLEAPKAPGVGLVDDGGCCGVCAKOL 60
 DB 1 MSSSTFRLAVAVTLHLTRIALSTCPACHCPLEAPKAPGVGLVDDGGCCGVCAKOL 60

OY 61 NEDSKTOPCDHTGKLECNFGASSTALKGICRAOSEGRPCENSRITYONGESFQPNCOHQ 120
 DB 61 NEDSKTOPCDHTGKLECNFGASSTALKGICRAOSEGRPCENSRITYONGESFQPNCHQ 120

OY 121 CTCIDGAVGCIPLCPQELSLPNLCGPNRLVKVTGQCCSEWVCDESDIKDPMEDDGLLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNLCGPNRLVKVTGQCCSEWVCDESDIKDPMEDDGLLG 180

OY 121 CTCIDGAVGCIPLCPQELSLPNLCGPNRLVKVTGQCCSEWVCDESDIKDPMEDDGLLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPNLCGPNRLVKVTGQCCSEWVCDESDIKDPMEDDGLLG 180

OY 181 KELGFDSEVBLTRNNELIANGKRSILKRLPFVFGMERILYNPL--GQOKIVQTTSSQ 238
 DB 179 --LGIDASEVELTRNNELIANGKSSILKRLPFVFGMERILYNPLAHGQKIVQTTSSQ 236

OY 239 CSKTCGAGISIRVNDPECELVKTRICEVRPCGOPYSSLKKGKSKTKSPPEYRF 298
 DB 237 CSKTCGAGISIRVNDPECELVKTRICEVRPCGOPYSSLKKGKSKTKSPPEYRF 296

OY 299 TYAGCLSVKKYRPYCGSCVDGRCTPOLTRTVAMRRCDEGTFSKVMYIOGCKNYN 358
 DB 297 TYAGCSSVKYRRPYCGSCVDGRCTPOLTRTVAMRRCDEGTFSKVMYIOGCKNYN 356

OY 359 CPANEAPFPYRLFNDIHKFRD 381
 DB 357 CPHPNEASFRLYSLFNDIHKFRD 379

RESULT 3
 CYR6_RAT STANDARD; PRT; 379 AA.
 ID CYR6_RAT
 AC Q9ES72;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE CYR61 protein precursor (Cysteine-rich, angiogenic inducer, 61)
 DE (Insulin-like growth factor-binding protein 10).
 GN CYR61 OR IGFBP10.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=20435857; PubMed=10852911;
 RA Albrecht C., von Der Kammer H., Mayhaus M., Klaudiny J., Schweizer M.,
 RT Nitsch R.M.;
 RT "Muscarinic acetylcholine receptors induce the expression of the
 RT immediate early growth regulatory gene Cyr61".
 RL J. Biol. Chem. 275:28929-28936(2000).
 CC -1- FUNCTION: MAY ACT AS ONE OF THE MANY GROWTH FACTOR-BINDING
 CC PROTEINS; PROMOTES PROLIFERATION, MIGRATION AND ADHESION (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 WFPC domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF218568; AAC14964.1; -;
 DR InterPro; IPR006208; Cys_knot.
 DR InterPro; IPR006207; Cys_knot.C.
 DR InterPro; IPR000867; Ins1_gro_fac-pr.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR001007; WFC_C.
 DR Pfam; PF00007; Cys_knot; 1.

DR Pfam: PF00219; IGFBP, 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR Pfam: PF00093; vwc; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; WVC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF BINDING; 1.
 DR PROSITE; PS50092; TSP1; 1.
 DR PROSITE; PS01208; WVC_1; 1.
 DR PROSITE; PS50184; WVC_2; 1.
 DR Growth factor binding; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 379
 FT DOMAIN 26 97
 FT DOMAIN 98 164
 FT DOMAIN 226 271
 FT DOMAIN 284 358
 FT DISULFID 284 321
 FT DISULFID 301 335
 FT DISULFID 312 351
 FT DISULFID 315 353
 FT DISULFID 320 357
 SQ SEQUENCE 379 AA; 41687 MW; 62BF0BBA4C5AFDE9 CRC64;

Query Match Best Local Similarity 91.2%; Score 1929; DB 1; Length 379;
 Matches 348; Conservative 10; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSSRIARALAVVTLHLTRLALSTCPACHCPLPAPKAPGVGLVRDCCGCCVYCAKOL 60
 DB 1 MSSRTKLAVAVTLHLTRLALSTCPACHCPLPAPKAPGVGLVRDCCGCCVYCAKOL 60
 QY 61 NEDCSKTOPCDHTKGLGECNFGASSALKKICRAQSGRCENSRITYONGESFOPNCKHQ 120
 DB 61 NEDCSKTOPCDHTKGLGECNFGASSALKKICRAQSGRCENSRITYONGESFOPNCKHQ 120
 QY 121 CTCIDGAVGCIPLCPQELSLPMLGCPNPLRVKVGQCCCEWVCDDESDINDPEDDGLG 180
 DB 121 CTCIDGAVGCIPLCPQELSLPMLGCPNPLRVKVGQCCCEWVCDDESDINDPEDDGLG 180
 QY 121 CTCIDGAVGCIPLCPQELSLPMLGCPNPLRVKVGQCCCEWVCDDESDINDPEDDGLG 178
 DB 121 CTCIDGAVGCIPLCPQELSLPMLGCPNPLRVKVGQCCCEWVCDDESDINDPEDDGLG 178
 QY 181 KLGFDASEVELTRNNELIAGKGRSLKRLPVFGMEPRILYNPL--QGCKLIVQTTSMQ 238
 DB 179 --LGFDASEVELTRNNELIAGKGRSLKRLPVFGMEPRILYNPL--QGCKLIVQTTSMQ 236
 QY 239 GSKTGTGISTRTVNDNPNCRVYKTRICEVRPCQPVYSSLAKKKCKSKTKSPVPVF 298
 DB 237 CSKSGGTGISTRTVNDNPNCRVYKTRICEVRPCQPVYSSLAKKKCKSKTKSPVPVF 296
 QY 239 TYAGCLSVKRYKRYCGSCVDGRCTPOLTRTKKRFRCDEDETFSKNVMIOQCKCNV 358
 DB 237 TYAGCSSVKRYKRYKRYCGSCVDGRCTPOLTRTKKRFRCDEDETFSKNVMIOQCKCNV 356
 QY 359 CPHANEAAPFYRLFNDFIHKFRD 381
 DB 357 CPHPNEASFRLYSLFNDFIHKFRD 379

RESULT 4
 CE10_CHICK STANDARD; PRT; 375 AA.
 AC P19336;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE CEF-10 protein precursor.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;

[1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=89145206; PubMed=2537491;
 RX Simmons D.L., Levy D.B., Yannoni Y., Erikson R.L.,
 RA Identification of a photobol ester-repressible v-src-inducible gene.;
 RT Proc. Natl. Acad. Sci. U.S.A. 86:1178-1182(1989).
 CC -1- FUNCTION: PROBABLE SECRETED REGULATORY PROTEIN.
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- INDUCTION: BY V-SRC.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 WVC domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: J04486; AAA8661.1; -
 DR PIR: A41428; A41428.
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot.C.
 DR InterPro: IPR000867; Ins1_gro_fac_pr.
 DR InterPro: IPR001007; WVC_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00219; IGFBP; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR Pfam: PF00093; vwc; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; WVC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF BINDING; 1.
 DR PROSITE; PS50092; TSP1; 1.
 DR PROSITE; PS01208; WVC_1; 1.
 DR PROSITE; PS50184; WVC_2; 1.
 DR Growth factor binding; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 375
 FT DOMAIN 98 164
 FT DOMAIN 223 268
 FT DOMAIN 281 355
 FT DISULFID 281 318
 FT DISULFID 298 332
 FT DISULFID 309 348
 FT DISULFID 312 350
 FT DISULFID 317 354
 SQ SEQUENCE 375 AA; 40651 MW; 95F28533BE3505AE CRC64;

Query Match Best Local Similarity 80.1%; Score 1695; DB 1; Length 375;
 Matches 311; Conservative 20; Mismatches 41; Indels 12; Gaps 7;

QY 1 MSSRIARALAVVTLHLTRLALSTCPACHCPLPAPKAPGVGLVRDCCGCCVYCAKOL 59
 DB 1 MSGAGARP-ALAAALICLARIALGSPCPAVCCCPAAAPQACAGVGLVDPGCCCCVYCAKOL 59
 QY 60 LNEDCSKTOPCDHTKGLGECNFGASSALKKICRAQSGRCENSRITYONGESFOPNCKHQ 119
 DB 60 LNEDCSKTOPCDHTKGLGECNFGASSALKKICRAQSGRCENSRITYONGESFOPNCKHQ 119
 QY 120 OCTCIDGAVGCIPLCPQELSLPMLGCPNPLRVKVGQCCCEWVCDDESDINDPEDDGLG 179
 DB 120 OCTCIDGAVGCIPLCPQELSLPMLGCPNPLRVKVGQCCCEWVCDDESDINDPEDDGLG 177
 QY 180 GKLGFDASEVELTRNNELIAGKGRSLKRLPVFGMEP--RLYNPLQGGCKLIVQTTSMQ 237

```

Db 178 SKERGLDASEGELTRNNELIAIVKG-GLKMLPVGSEPSORAFENP-----KCIYQTSMS 232
QY 238 OCSKTGCTGISTRTVNDNPECRILVKETRICIEVRRCGQPVYSLKKGKCKSKTKSPSPVR 297
Db 233 QCSKTGCTGISTRTVNDNPECRILVKETRICIEVRRCGQPVYSLKKGKCKSKTKSPSPVR 292
QY 298 FTYAGCLSVKRYPRKYGSCVDGRCTPOLTRIVKMFRCDEGTFSKNVMIOSCKNY 357
Db 293 FTYAGCSSLVKRYPRKYGSCVDGRCTPOLTRIVKMFRCDEGTFSKNVMIOSCKNY 352
QY 358 NCPHANEAPFYYRLFNDIHKFRD 381
Db 353 NCPHANEAPFYYRLVNDIHKFRD 375

RESULT 5
CTGF_MOUSE STANDARD; PRT; 347 AA.
ID Q9RIE9; Q9WVS1;
AC Q9RIE9; Q9WVS1;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Connective tissue growth factor precursor.
GN CTGF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP MEDLINE=20145935; PubMed=10679821;
RX Xu J., Smock S.L., Safadi F., Rosenzweig A.B., Odgren P.R.,
RA Marks S.C. Jr., Owen T.A., Popoff S.N.:
RT "Cloning the full-length cDNA for rat connective tissue growth factor:
RT implications for skeletal development."
RT J. Cell. Biochem. 77:103-115(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Tezuka K., Tamatani T.:
RT "Rattus norvegicus connective tissue growth factor."
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATPACTANT SECRETED BY
CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL
CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
CC SYNTHESIS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WPC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF120275; BAB39132.1; -
CC EMBL; AB023068; BAB82125.1; -
CC InterPro: IPR006207; Cys_knot.
CC InterPro: IPR006207; Cys_knot_C.
CC InterPro: IPR000867; Ins1_gro_fac_pr.
CC InterPro: IPR000884; TSP1.
CC InterPro: IPR001007; WVF_C.
CC Pfam: PF00007; Cys_knot; 1.
CC Pfam: PF00219; IGFBP; 1.
CC Pfam: PF00090; TSP_1; 1.
CC Pfam: PF00093; WVC; 1.
CC SMART; SM00041; CT; 1.

```

```

DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; WVC; 1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS00222; IGFBINDING; 1.
DR PROSITE; PS00092; TSP1; 1.
DR PROSITE; PS01208; WVF_C; 1.
DR PROSITE; PS01084; WVF_C_2; 1.
DR Cell adhesion; DNA synthesis; Extracellular matrix; Signal.
KW SIGNAL
FT CHAIN 1 24
FT DOMAIN 25 347
FT DOMAIN 31 98
FT DOMAIN 99 165
FT DOMAIN 196 241
FT DOMAIN 234 328
FT DISULFID 254 291
FT DISULFID 271 305
FT DISULFID 282 321
FT DISULFID 285 323
FT DISULFID 290 327
FT CONFLICT 35
FT CONFLICT 94
SQ SEQUENCE 347 AA; 37756 MW; CFBETAL9766B7B16 CRC64;

Query Match 45.18; Score 955; DB 1; Length 347;
Best Local Similarity 46.28; Pred. No. 2.2e-62;
Matches 176; Conservative 58; Mismatches 107; Indels 40; Gaps 7;

QY 1 MSSRIARALALVYTLTLRLAL-STCPACHQPLE-APKCAPGVGLVRDCCGCKYCAK 58
Db 1 MASVAGPVSIALVLLCTRPATGQDCSAGQCAAEAPRCAPGVSLVLDGCGCCRCYCAK 60
QY 59 QLNEDCSKTOPCHTKGLENGFASSTALGICRAOSEGPCEYNSHTYNGSFQPCNC 118
Db 61 QLSCLTERPDCPHKGLFCDFGSPANRKGICVCTAK-DGAPCVFGSGSVYSGESFSSCK 119
QY 119 HOCCTIDGAVGCIPLCPQELSPNLGCPNRLKYNVGGCCCEWVCDEDSIKDMEDQDL 178
Db 120 YQCTCLDGAAGVCPVLCMDVRLPSDCPPFRVRVYLPKCCCEWVCDEP----- 167
QY 179 LGRLEFGDASEVELTRNNELIAVKGKSLRL-PVQMEPRILYNPLQGGKCIYQTSMS 237
Db 168-----KDTIVYGPAALARLDTRGPDPTM-----RANCLVQTSMS 205
QY 238 QCSKTGCTGISTRTVNDNPECRILVKETRICIEVRRCGQPVYSLKKGKCKSKTKSPSPVR 297
Db 206 ACSKTGCMGISTRTVNDNPECRILVKETRICIEVRRCGQPVYSLKKGKCKSKTKSPSPVR 265
QY 298 FTYAGCLSVKRYPRKYGSCVDGRCTPOLTRIVKMFRCDEGTFSKNVMIOSCKNY 357
Db 266 FELSGLTSVYTRAKRGVCTDGRCCPHRTTLVPEFKCPDDELKKNMFMFKTACHY 325
QY 358 NCPHANEAPFYY--RLFNDI 376
Db 326 NCPGDMIDFESLTKYKMGDM 346

RESULT 6
CTGF_MOUSE STANDARD; PRT; 348 AA.
ID P29268; Q922U0;
AC P29268; Q922U0;
DT 01-DEC-1992 (Rel. 24, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Connective tissue growth factor precursor (FISP-12 protein)
DE (Hypertrophic chondrocyte-specific protein 24).
GN CTGF OR FISP12 OR FISP-12 OR HCS24.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

RP SEQUENCE FROM N.A.
 RX MEDLINE-91363290; PubMed-1886698;
 RA Rybeck R.-P., Macdonald-Bravo H., Mattei M.-G., Bravo R.;
 RT "Structure, mapping, and expression of flsp-12, a growth factor-
 RL inducible gene encoding a secreted cysteine-rich protein.";
 RN Cell Growth Differ. 2:225-233(1991).
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91229699; PubMed-2029337;
 RA Brunner A., Chinn J., Neubauer M.G., Purchio A.F.;
 RT "Identification of a gene family regulated by transforming growth
 RL factor-beta.";
 RN DNA Cell Biol. 10:293-300(1991).
 [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-22388257; PubMed-12477932;
 RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Morina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carroll P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratnam P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzyzinski M.I., Skalska U., Smallus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RL human and mouse cDNA sequences.";
 RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [4]
 RP FUNCTION, AND SUBCELLULAR LOCATION.
 RX MEDLINE-97327410; PubMed-9184077;
 RA Kireeva M.L., Latifkic B.V., Kolesnikova T.V., Chen C.C., Yang G.P.,
 RA Ahler A.S., Lau L.F.;
 RT "Cyt61 and flsp12 are both ECM-associated signalling molecules:
 RL activities, metabolism, and localization during development.";
 RN Exp. Cell Res. 233:63-77(1997).
 [5]
 RP FUNCTION.
 RX MEDLINE-99182484; PubMed-10082563;
 RA Babic A.M., Chen C.C., Lau L.F.;
 RT "flsp12/mouse connective tissue growth factor mediates endothelial
 RT cell adhesion and migration through integrin alphavbeta3, promotes
 RT endothelial cell survival, and induces angiogenesis in vivo.";
 RN Mol. Cell. Biol. 19:2958-2966(1999).
 CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
 CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
 CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL
 CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
 CC SYNTHESIS (BY SIMILARITY).
 CC -1- SUBUNIT: Monomer (By similarity).
 CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
 CC SOLUBLE FORM.
 CC -1- TISSUE SPECIFICITY: TESTIS, SPLEEN, KIDNEY, LUNG, HEART, AND BRAIN
 CC (LOWEST LEVEL IN TESTIS AND HIGHEST IN LUNG).
 CC -1- INDUCTION: By growth factors.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 VWF domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entries requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M70641; AAA37627.1; -
 CC EMBL: M70642; AAA37628.1; -
 CC EMBL: M80263; AAA73135.1; -
 CC EMBL: BC006783; AAH06783.1; -
 CC PIR: A40578; A40578.
 CC MGI: 95537; Ctgf.
 CC GO: GO:0005578; C:extracellular matrix; IDA.
 CC GO: GO:0008202; F:heparin binding activity; IDA.
 CC GO: GO:0005178; F:integrin binding activity; IDA.
 CC GO: GO:0001525; P:angiogenesis; IDA.
 CC GO: GO:0016477; P:cell migration; IDA.
 CC GO: GO:0007160; P:cell-matrix adhesion; IDA.
 CC GO: GO:0008543; P:FGF receptor signaling pathway; IDA.
 CC GO: GO:0007228; P:integrin-mediated signaling pathway; IDA.
 CC GO: GO:0001503; P:osification; IMP.
 CC InterPro: IPR006208; Cys_knot.
 CC InterPro: IPR006207; Cys_knot_C.
 CC InterPro: IPR000867; Insl_gro_fac_pr.
 CC InterPro: IPR000884; TSP1.
 CC InterPro: IPR001007; VWF_C.
 CC Pfam: PF00007; Cys_knot; 1.
 CC Pfam: PF00219; IGFBP; 1.
 CC Pfam: PF00090; TSP1; 1.
 CC Pfam: PF00093; vwc; 1.
 CC SMART: SM00041; CT; 1.
 CC SMART: SM00121; IB; 1.
 CC SMART: SM00209; TSP1; 1.
 CC SMART: SM00214; VWC; 1.
 CC PROSITE: PS01185; CTCK; 1.
 CC PROSITE: PS01225; CTCK_2; 1.
 CC PROSITE: PS00222; IGF_BINDING; 1.
 CC PROSITE: PS50092; TSP1; 1.
 CC PROSITE: PS01208; VWF_C; 1.
 CC PROSITE: PS01084; VWF_C_2; 1.
 CC Cell adhesion; DNA synthesis; Extracellular matrix; signal.
 CC POTENTIAL.
 CC CONNECTIVE TISSUE GROWTH FACTOR.
 CC IGFBP.
 CC VWF.
 CC TSP TYPE-1.
 CC CTCK.
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC BY SIMILARITY.
 CC E -> K (IN REF. 1).
 CC SEQUENCE 348 AA; 37794 MW; 4D7B6D9089174049 CRC64;
 SQ
 Query Match 45.0%; Score 952.5; DB 1; Length 348;
 Best Local Similarity 46.6%; Pred. No. 3.3e-62;
 Matches 178; Conservative 57; Mismatches 106; Indels 41; Gaps 8;
 QY 1 MSSRIARLAVLTVTLHL-TRLAL-STCPACHCLE-APKCAPGVGLVRDGGCCCKCA 57
 DB 1 MIAVAGPISLALVLLALCTPRATGQDCSACQCAEAAPHCAPGVSIVLDGCGCRCA 60
 QY 58 KQLNEDCKPTPCDHTKGLCECNFGASSTALNGICRAQSGRCPCEYNSRIYNGESFPQNC 117
 DB 61 KQGLGELCEPDCPDHKLFDGSPANRKIGVCTAK-DGAPCVGSGSVYSGESFQSSC 119
 QY 118 QHQCCTIDGAVGICPLCPQELSLPVLGCPNRLKAVYTGCCCEBWCDEDSIKDPNEDDG 177
 DB 120 KYQCCLDGAAGVCPVLCMDVRLPSPDCPFRFRVRLPKCCBEMWCDEP----- 168
 QY 178 LLGRKLGDASVELTRNNELIANGKSKRL-LPVFMEPRILYNPLQGGKCIQVOTSM 236
 DB 169 -----KDRIVGPAALAAVRLERTFPPDPMM-----RANCLVQTTW 205
 QY 237 SCSKTCGTGISTRTVNDNPECRVLKETRICEVPRPCGPVYSLSLKGKCKSKTKSPBPV 296

```

Db      206  SACSCTGCGISIRVYNDMTFCLERKQSLCWMRCPCEADLEENIKKGCIRPKIAXKV 265
QY      297  RFTYACCLSVKTRPRYCCSCVDGRCCTQTLRTVMRRCEDEFFSNVMATQSKCN 356
Db      266  KEPLSCTGCTVKYKRAFCVCCTDGRCTPRTTTLVEFKCPDGEIKMKNMFIKTCACH 325
QY      357  YNCPHANEAAFPFY--RLFNDI 376
Db      326  YNCPGNDIPESTLYTKMTGDM 347

RESULT 7
CTGF_HUMAN STANDARD; PRT; 349 AA.
ID      P29279; 0960X2;
DT      01-DEC-1992 (Rel. 24, Created)
DT      01-DEC-1992 (Rel. 24, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Connective tissue growth factor precursor (Hypertrophic chondrocyte-
DE      specific protein 24).
GN      CTGF OR HCS24.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE-Umbilical vein endothelial cells;
RX      MEDLINE=91373462; PubMed=1654338;
RA      Bradham D.M., Igarashi A., Potter R.L., Grotendorst G.R.;
RT      "Connective tissue growth factor: a cysteine-rich mitogen secreted by
RT      human vascular endothelial cells is related to the SRC-induced
RT      immediate early gene product CEF-10."
RL      J. Cell Biol. 114:1285-1294(1991).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE-Umbilical vein endothelial cells;
RX      MEDLINE=93187114; PubMed=1293144;
RA      Igarashi A., Bradham D.M., Okochi H., Grotendorst G.R.;
RT      "Connective tissue growth factor."
RL      J. Dermatol. 19:642-643(1992).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=97207446; PubMed=9054739;
RA      Oemar B.S., Werner A., Garner J.M., Do D.D., Godoy N., Nauck M.,
RA      Marz W., Rupp J., Pech M., Luescher T.F.;
RT      "Human connective tissue growth factor is expressed in advanced
RT      atherosclerotic lesions."
RL      Circulation 95:831-839(1997).
RN      [4]
RP      SEQUENCE FROM N.A.
RA      Cobley V.;
RL      Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN      [5]
RP      FUNCTION.
RC      TISSUE-Chondrocytes;
RX      MEDLINE=20080284; PubMed=10614647;
RA      Nakaiishi T., Nishida T., Shimo T., Kobayashi K., Kubo T.,
RA      Tametani T., Tezuka K., Takigawa M.;
RT      "Effects of CTGF/HCS24, a product of a hypertrophic chondrocyte-
RT      specific gene, on the proliferation and differentiation of
RT      chondrocytes in culture."
RL      Endocrinology 141:264-273(2000).
RN      [6]
RP      FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
RP      VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
RP      DIFFERENTIATION OF CHONDROCYTES. MEDIATES CELL ADHESION AND
RP      ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA SYNTHESIS.
RN      [7]
RP      SUBUNIT: Monomer
RN      [8]
RP      SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
RN      [9]
RP      SOLUBLE FORM (BY SIMILARITY).
RN      [10]
RP      ALTERNATIVE PRODUCTS:
RN      [11]
RP      Event-Alternative splicing; Named isoforms=2;

```

```

CC      Name-Long;
CC      IsoId=P29279-1; Sequence=Displayed;
CC      Name-Short;
CC      IsoId=P29279-2; Sequence=VSP_002460;
CC      Note=No experimental confirmation available;
CC      -1- SIMILARITY: Contains 1 IGFBP domain.
CC      -1- SIMILARITY: Contains 1 WFCC domain.
CC      -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC      -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: M92934; AAA91279.1; -
DR      EMBL: X78947; CAA55544.1; -
DR      EMBL: AL354866; CAC44023.1; -
DR      PIR: A40551; A40551.
DR      Genew: HGNC:2500; CTGF.
DR      MIM: 121009; -
DR      GO: GO:0005578; C:extracellular matrix; TAS.
DR      GO: GO:0005886; C:plasma membrane; TAS.
DR      GO: GO:0005520; F:insulin-like growth factor binding activity; TAS.
DR      GO: GO:0008151; P:cell growth and/or maintenance; TAS.
DR      GO: GO:0008544; P:epidermal differentiation; TAS.
DR      GO: GO:0009611; P:response to wounding; TAS.
DR      InterPro: IPR006208; Cys_knot.
DR      InterPro: IPR006207; Cys_knot.
DR      InterPro: IPR000867; Ins1_gro_fac-pr.
DR      InterPro: IPR000884; TSP1.
DR      InterPro: IPR001007; WFCC.
DR      Pfam: PF00007; Cys_knot; 1.
DR      Pfam: PF00219; IGFBP; 1.
DR      Pfam: PF00090; TSP_1; 1.
DR      Pfam: PF00093; WVC; 1.
DR      SMART: SM00041; CT; 1.
DR      SMART: SM00121; IB; 1.
DR      SMART: SM00209; TSP1; 1.
DR      SMART: SM00214; WVC; 1.
DR      PROSITE: PS01185; CTCK_1; 1.
DR      PROSITE: PS01225; CTCK_2; 1.
DR      PROSITE: PS00222; IGF_BINDING; 1.
DR      PROSITE: PS00092; TSP1; 1.
DR      PROSITE: PS01208; WFCC_1; 1.
DR      PROSITE: PS0184; WFCC_2; 1.
KW      Cell adhesion; DNA synthesis; Extracellular matrix; Signal;
KW      Alternative splicing.
FT      SIGNAL 1..26
FT      CHAIN 27..349
FT      DOMAIN 33..100
FT      DOMAIN 101..167
FT      DOMAIN 198..243
FT      DOMAIN 256..330
FT      DISULFID 256..283
FT      DISULFID 273..307
FT      DISULFID 284..323
FT      DISULFID 287..325
FT      DISULFID 292..329
FT      CARBOHYD 28..28
FT      CARBOHYD 225..225
FT      CARBOHYD 225..225
FT      VARSPLIC 172..198
FT      CONFLICT 83..83
FT      SEQUENCE 349 AA; 38069 MW; 0ECF8470B357EA95 CRC64;
Query Match 44.8%; Score 947; DB 1; Length 349;
Best Local Similarity 46.6%; Pred. No. 8; 3e-62;
Matches 174; Conservative 58; Mismatches 101; Indels 40; Gaps 8;

```

```

QY 9 LALVLTLLHLRLAL-STCPACHCPLE-APKCAPGVGLVDRGGCCCKVCAROLNEDCSK 66
DB 11 VAFVLLALCSRPANVGSCGRCRDPAPRCAGVSLVLDGCGCCRCVAKOLGEICTE 70
QY 67 TOPCDHTKGLCNCNGASSTALKGICRAOSEGRPCENRITONGESFPNCOHOCTCIDG 126
DB 71 RDPDCHGGLGCLDSCSPANRRIGVCTAK-DGAPCITFGGTYSRGSFSSCKYOCCTCIDG 129
QY 127 AVGCIPICPOELSLPNLGCNPNRLVKYVGOCCEEWDEDESDIKDMEDDGLGKELGDA 186
DB 130 AVGCIPICMDYRLPSPCCPPRRYKLPKCCCEEWDCDEP-----KQ----- 172
QY 187 ASEVELTRNNELIANGKRSIKRL-PVFGMEPRILYNPLDQCKIVQTTSMQSKTCGT 245
DB 173 -----TVGPPALAYRLDEDFGDPPTMI-----RANCLVQTTSMQSKTCGM 215
QY 246 GISTRTVNDNPECRIVKTRICEVRPCGQPYSSLKKGKCKSKTKSPKPEPRFTYAGLS 305
DB 216 GISTRTVNDNNAKSKLEKOSRLCQVAPCEADLEENIKKCKIRPKISKPIKELSGCTS 275
QY 306 VKKYPKYGSCVDGRCTPOLTTRIVKMRFRCEDEGTFPSKVMVMIOSCKCNYNCPHANE 365
DB 276 MKYTRAKFCGCTDGRCTPRTTLLPVEFKCPDGEVKKMMMFICTGACHNCPGNDI 335
QY 366 AEPFY--RLFNDI 376
DB 336 FESLYRRMGYDM 348

```

RESULT 8

CTGF_BOVIN STANDARD: PRT: 349 AA.

```

AC 018739; Q99GL71;
DF 15-JUL-1998 (Rel. 36, Created)
DF 28-FEB-2003 (Rel. 41, Last sequence update)
DF 28-FEB-2003 (Rel. 41, Last annotation update)
DE Connective tissue growth factor precursor.
GN CTGF.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aorta;
RA Liljensiek B., Iain Z., Folsis T., Schlimanski M., Bierhaus A.,
RL Kautz M., Kaufmann G., Schweigert L., Ziegler R., Nawroth P.P.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Mathias M., Schwitters C., Hove M., Rupp S., Froman N.E.;
RT "Bovine connective tissue growth factor: organization of the
RT chromosomal gene and demonstration of promoter activity.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOATTRACTANT SECRETED BY
CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL
CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
CC SYNTHESIS (BY SIMILARITY).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
CC SOLUBLE FORM (BY SIMILARITY).
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WFPC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF000137; AAB6596.1; -
DR EMBL: AF309555; AAG30290.1; -
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot_C.
DR InterPro: IPR000867; Insl_gro_fac-pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; WFPC.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; TSP_1; 1.
DR Pfam: PF00093; WFC; 1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; WFC; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS00092; TSP1; 1.
DR PROSITE: PS01208; WFPC_1; 1.
DR PROSITE: PS0184; WFPC_2; 1.
DR Cell adhesion; DNA synthesis; Extracellular matrix; Signal.
KW Cell adhesion; DNA synthesis; Extracellular matrix; Signal.
FT SIGNAL 1 26
FT CHAIN 27 349
FT DOMAIN 33 100
FT DOMAIN 101 167
FT DOMAIN 198 243
FT DOMAIN 256 330
FT DISULFID 256 293
FT DISULFID 273 307
FT DISULFID 287 325
FT DISULFID 292 329
FT CONFLICT 28 28
FT CONFLICT 58 58
FT CONFLICT 88 88
FT CONFLICT 103 104
FT CONFLICT 166 169
FT CONFLICT 184 184
FT CONFLICT 200 200
FT CONFLICT 209 209
FT CONFLICT 269 269
FT CONFLICT 284 284
SQ SEQUENCE 349 AA; 37924 MW; 5FEC8EB83EFBAF99 CRC64;
Query Match 44.1%; Score 933; DB 1; Length 349;
Best Local Similarity 46.4%; Fred. No. 8.5e-61;
Matches 172; Conservative 59; Mismatches 102; Indels 38; Gaps 8;
QY 10 ALVLTLLHLRLALS-TCPACHCPLE-APKCAPGVGLVDRGGCCCKVCAROLNEDCSK 67
DB 12 AVFVLLALCSRPANVGSCGRCRDPAPRCAGVSLVLDGCGCCRCVAKOLGEICTER 71
QY 68 QPCHHTKGLCNCNGASSTALKGICRAOSEGRPCENRITONGESFPNCOHOCTCIDGA 127
DB 72 DPCDPHKLGLDFDGSFANRRIGVCTAK-DGAPCVFGGTIVYSGSFSSCKYOCCTCIDGS 130
QY 128 VGCIPICPOELSLPNLGCNPNRLVKYVGOCCEEWDEDESDIKDMEDDGLGKELGDA 187
DB 131 VGCIPICVDVRLPSPCCPPRRYKLPKCCCEEWDCDEP-----PREHT--VVGPPALAYR 183
QY 188 SEVELTRNNELIANGKRSIKRL-PVFGMEPRILYNPLDQCKIVQTTSMQSKTCGTGI 247
DB 184 PE-----DTFGDPPTMI-----RANCLVQTTSMQSKTCGTGIGI 217
QY 248 STRVYNDNPECRIVKTRICEVRPCGQPYSSLKKGKCKSKTKSPKPEPRFTYAGLSVK 307
DB 218 STRVYNDNAPRLKSKLEKOSRLCQVAPCEADLEENIKKCKIRPKISKPIKELSGCTSMK 277
QY 308 KYRKYKSCVDGRCTPOLTTRIVKMRFRCEDEGTFPSKVMVMIOSCKCNYNCPHANEAF 367

```



```

Db 278 TYRACGCVCTDRCCTPHRTTLPVFPKCPDSEVWKKSMFKTCAChNCGGNDIEF 337
QY 368 PFY--RLFNDI 376
Db 338 SLTYRKMVGDM 348

RESULT 9
CTGF_PIG STANDARD; PRT; 349 AA.
AC 01913;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Connective tissue growth factor precursor.
GN CTGF.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RX MEDLINE=97390475; PubMed=9242708;
RA Brigstock D.R., Steffen C.L., Kim G.Y., Vegunta R.K., Diehl J.R.,
RA Harding P.A.;
RT "Purification and characterization of novel heparin-binding growth
RT factors in uterine secretory fluids. Identification as heparin-
RT regulated Mr 10,000 forms of connective tissue growth factor.";
RL J. Biol. Chem. 272:20275-20282(1997).
CC -1- FUNCTION: MAJOR CONNECTIVE TISSUE MITOGENETIC SECRETED BY
CC VASCULAR ENDOTHELIAL CELLS. PROMOTES PROLIFERATION AND
CC DIFFERENTIATION OF CHONDROCYTES (BY SIMILARITY). MEDIATES CELL
CC ADHESION AND ENHANCES FIBROBLAST GROWTH FACTOR-INDUCED DNA
CC SYNTHESIS (BY SIMILARITY).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: FOUND IN THE EXTRACELLULAR MATRIX AND AS A
CC SOLUBLE FORM (BY SIMILARITY).
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WFC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL, U83916; AAC48756.1; -.
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Ins_lgro_C.
DR InterPro: IPR000867; Ins_lgro_C.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; WFC_C.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; tsp_1; 1.
DR SMART: SM00041; Cys_1.
DR SMART: SM00121; IB_1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; WFC; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00232; IGF_BINDING; 1.
DR PROSITE: PS00092; TSP1; 1.
DR PROSITE: PS01208; WFC_1; 1.
DR PROSITE: PS01014; WFC_2; 1.
DR Cell adhesion; DNA synthesis; Extracellular matrix; Signal.

```

```

FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 349 CONNECTIVE TISSUE GROWTH FACTOR.
FT DOMAIN 33 100 IGFBP.
FT DOMAIN 101 167 WFC.
FT DOMAIN 198 243 TSP TYPE-1.
FT DOMAIN 256 330 CTCK.
FT DISULFID 256 293 BY SIMILARITY.
FT DISULFID 273 307 BY SIMILARITY.
FT DISULFID 284 323 BY SIMILARITY.
FT DISULFID 287 325 BY SIMILARITY.
FT DISULFID 292 329 BY SIMILARITY.
SQ SEQUENCE 349 AA; 38007 MW; BB510E2B2B52D4A0 CRC64;

Query Match 42.8%; Score 906; DB 1; Length 349;
Best Local Similarity 44.9%; Pred. No. 7.7e-59;
Matches 167; Conservative 57; Mismatches 108; Indels 40; Gaps 7;

QY 10 ALVYTLHLRLAL-STCPAACHCP-LEAPKCAPGVGLYRDGGCCCKVCAKQINEDCSKT 67
Db 12 AFVLLALCSRPASGDCSGGCCAGKRRACAGSYLVLDGCGCCRLCAKQIGELCTER 71
QY 68 QPCDHTKGLGECNFGASSIALKIGICRAOSEGRPCVNSRIYQNGESFQPMQCQHOCTCIDA 127
Db 72 DPCDPKRGJLFCDFGSPANKRIGVCTAK-DGAPCVFGITVRSSEFSQSKYQCTCIDGA 130
QY 128 VGCIPICPQELSPNIGCENPRVYKVTGGCCCEWVDEDSINDPMEDOGILGKELGFA 187
Db 131 VGVVPLCSMDVRLPDCPPRRVKLPKGCCEWVDEP----- 169
QY 188 SEVELRNNELIANGGRSLKRL-PVFGMEPRILYPLQKCIYVTTSMQSKTCGTG 246
Db 170 -----KDHVYVAGALAYRIEDFFGDPIM-----RANLVYGTTERSACSKTCGNG 216
QY 247 ISTRVNDNDECHLVKETRICEVRPGQPYVSLKGGKSKSKSPKPEVRYAGCLSV 306
Db 217 ISTRVYNDNAFCLEKOSRLCWVRPEADLEENIKKRCIRTPKISKPKVELSGCTSV 276
QY 307 KRYRPPYGCYDVRCPQLPRTYKMRRCGECEFFSNVMMIOGCKNYPHANEAA 366
Db 277 KTYRAFGCVCTDRCCTPHRTTLPVFPKCPDSEVWKKSMFKTCAChNCGGNDIEF 336
QY 367 PFY--RLFNDI 376
Db 337 SLTYRKMVGDM 348

RESULT 10
NOV_CHICK STANDARD; PRT; 351 AA.
ID NOV_CHICK
AC P28686;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE NOV protein precursor (Nephroblastoma overexpressed gene protein).
GN NOV.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Brown Leghorn;
RX MEDLINE=92107157; PubMed=1309586;
RA Joliet V., Martinerie C., Dambrine G., Plassiat G., Brissac M.,
RA Crochet J., Perbal B.;
RT "Proximal rearrangements and overexpression of a new cellular gene
RT (nov) in myeloblastosis-associated virus type 1-induced
RT nephroblastomas";
RL Mol. Cell. Biol. 12:10-21(1992).
CC -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
CC GROWTH REGULATION. ITS OVEREXPRESSION IS ASSOCIATED WITH
CC TUMORIGENESIS AND EXPRESSION OF A N-TERMINAL-TRUNCATED VERSION

```

CC OF NOV GENE IN CHICKEN EMBRYONIC FIBROBLASTS (CEF) IS SUFFICIENT
 CC TO INDUCE THE TRANSFORMATION OF CEF IN VITRO.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: BRAIN AND HEART, AND AT A LOWER LEVEL IN
 CC MUSCLE AND INTESTINE, IN THE EMBRYO, LUNG AND LESS SO IN BRAIN AND
 CC SPLEEN, IN ADULT CHICKEN.
 CC -1- DEVELOPMENTAL STAGE: MAV1-INDUCED NEPHROBLASTOMAS EXPRESS A HIGH
 CC LEVEL OF NOV GENE WHOSE TRANSCRIPTION IS NORMALLY ARRESTED IN
 CC ADULT KIDNEY.
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 WFMC domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X59284; CAA41975.1; -
 DR PIR: S20078; S20078.
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Ins1_gro_fac-pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; WFMC_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00219; IGFBP; 1.
 DR Pfam: PF00090; TSP_1; 1.
 DR Pfam: PF00093; WVC; 1.
 DR Pfam: PF00041; CT; 1.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00209; TSP1; 1.
 DR SMART: SM00214; WVC; 1.
 DR PROSITE: PS01185; CTCK_1; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR PROSITE: PS00222; IGF_BINDING; 1.
 DR PROSITE: PS00092; TSP1; 1.
 DR PROSITE: PS01208; WFMC_1; 1.
 DR PROSITE: PS0184; WFMC_2; 1.
 DR PROTO-ONCOGENE; Growth factor; Signal.
 KW SIGNAL 1 24
 FT CHAIN 25 351
 FT DOMAIN 31 103
 FT DOMAIN 104 170
 FT DOMAIN 201 246
 FT DOMAIN 258 332
 FT DISULFID 258 295
 FT DISULFID 275 309
 FT DISULFID 286 325
 FT DISULFID 289 327
 FT DISULFID 294 331
 FT CARBOHYD 274 274
 SQ SEQUENCE 351 AA; 38268 MW; 1ECB3FA30586C797 CRC64;
 Query Match 40.2%; Score 851.5; DB 1; Length 351;
 Best Local Similarity 44.0%; Pred. No. 6,9e-55;
 Matches 164; Conservative 44; Mismatches 104; Indels 61; Gaps 8;
 QY 11 LVVTLHLRLALS---TCPAAC--HCPLEAPKAPGVGLVDRGCGCCCKCAKOLNEDC 64
 DB 12 LLLLLLLRPCEVSGREAACPFGCGRCPAEPFPAVLDGCGCLVCARORGESC 71
 QY 65 STQPCDHTKGLKCFAGSSTLAKICRAQSGRCCEVNSRYONGESFQPCQHCQCCI 124
 DB 72 SPLPCDDSSGGLYCDRGGEDGGAGICAW-LEGDMCVDGMITYRNGETFPQCKYQCTCR 130
 QY 125 DGAAGCIPLCQPELSPNIGCPNPRVRYVKGCCCEWYCCDEDSIDDPEDMDGGLG---- 180
 DB 131 DGQIGCLPRCNIIGLILPGDPDCEFPKRIEYRGCCCKRWVC-----DPRDEV--LLGGRAM 182

QY 181 -----KELGPDASEVELLTNNELIANGKGRSLKRLPFGMERILLYNPLQOKCIYQTT 234
 DB 183 AAYROEATLIDVSD-----SANCIJEQTT 207
 QY 235 SMSGCKTGCGTGSTRVTNNDNPECRLYKETRICVPRPGQPVYSLKGRKSKTKKSP 294
 DB 208 EWSACSCKSGMGFSTRYTNRQCEMYKQTRLCMRCEBDEPSD-KKAKKCIQTKSMK 266
 QY 295 PVRETYAGCLSVKKRYRKYRGSCVDRCCPQULTRYKMRFCDEGETFSKNVMIQSC 354
 DB 267 AVREYKNCISVYQYKRYRGGLCNDGRCCPDPHNTKITQVERPCQGFLLKPPMLINTCV 326
 QY 355 CNINCPHANEAF 367
 DB 327 CHGNCPOSNNAF 339
 RESULT 11
 NOV_COTUA STANDARD; PRT; 353 AA.
 AC P42642;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NOV protein precursor (Nephroblastoma overexpressed gene protein).
 GN NOV.
 OS Coturnix coturnix japonica (Japanese quail).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Coturnix
 CC NCBI_Taxid=93934;
 RN [1]
 RP Weiskirchen R., Bister K.;
 RA Submitted (AUG-1994) to the EMBL/Genbank/DBJ databases.
 RL -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
 CC GROWTH REGULATION (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- SIMILARITY: Contains 1 IGFBP domain.
 CC -1- SIMILARITY: Contains 1 WFMC domain.
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
 CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U13063; AAA21128.1; -
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Ins1_gro_fac-pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; WFMC_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00219; IGFBP; 1.
 DR Pfam: PF00090; TSP_1; 1.
 DR Pfam: PF00093; WVC; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00121; IB; 1.
 DR SMART: SM00209; TSP1; 1.
 DR SMART: SM00214; WVC; 1.
 DR PROSITE: PS01185; CTCK_1; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR PROSITE: PS00222; IGF_BINDING; 1.
 DR PROSITE: PS00092; TSP1; 1.
 DR PROSITE: PS01208; WFMC_1; 1.
 DR PROSITE: PS0184; WFMC_2; 1.
 DR PROTO-ONCOGENE; Growth factor; Signal.
 KW

```

FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 353 NOV PROTEIN.
FT DOMAIN 33 105 IGFBP.
FT DOMAIN 106 172 VMFC.
FT DOMAIN 203 248 TSP TYPE-1.
FT DOMAIN 260 334 CTCK.
FT DISULFID 260 297 BY SIMILARITY.
FT DISULFID 277 311 BY SIMILARITY.
FT DISULFID 288 327 BY SIMILARITY.
FT DISULFID 291 329 BY SIMILARITY.
FT DISULFID 296 333 BY SIMILARITY.
FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 353 AA; 38667 MW; 717D9F853382EB9 CRC64;

Query Match 40.1%; Score 848; DB 1; Length 353;
Best Local Similarity 43.3%; Pred. No. 1,2e-54;
Matches 164; Conservative 45; Mismatches 104; Indels 66; Gaps 8;

QY 10 ALVVTLLHTRAL-----STCPAAC--HCPLEAPKAPGVGLVRDGCCKVKCAR 58
DB 8 SLVYVLLLLLLLLLRLPSEVNGREARPCRCGGRCPAEPRAVAVLDGCGCLVCAR 67
QY 59 QLNEDCSKTPCHRTGKLECNFASSTALKGICRAOSEGRPCERYNSRTYONGSPQNCQ 118
DB 68 QRESECSPLPCDESGGLYCDRGPEDGGGTGICMV--LEGDNVYFGDMITVNGETFOPSCK 126
QY 119 HOCTCIDGAVGCIPLCPQELSLNLCGPNRLKVTGQCCCEWVCDSDTKDMEDODGL 178
DB 127 YQCTCRDQIGLCPLRCMLGLLGPDPCEPRKTEVPECEKWC-----EPREDEV--L 178
QY 179 LG-----KELGFDASEVELTRNNELIAGVGRSLRLVFGMEPRILYNPLQGR 228
DB 179 LGGFAMAAYRQENATLGDVSD-----SSAN 203
QY 229 CIYQTTSMQSCSTCTGCGISTRYTNDNPECRLYKERRICVRPGOGPVYSLKKKKCSK 288
DB 204 CIGQTEWMSACSSCGSGFSTRYTNRQOCENWKQTRLCMMARCENEPSPD--KKGKCIIR 262
QY 289 TKKSPREFTYAGCLSVKKYRPRKCGSCVDGRCPCPOLRTYKMRPRCDGDFESKNV 348
DB 263 TKKSMKAVRFRFYKNCSTVQYTKPRYGLCLMDGRCCKPHNTKTIQVEFRCPQGRFLKPKPM 322
QY 349 MIQSCKCNVNCPRHANAFAF 367
DB 323 LINTCVCHGNCPOSNNAAFE 341

RESULT 12
NOV_HUMAN STANDARD; PRT; 357 AA.
ID NOV_HUMAN
AC P48745; Q96BY5;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE NOV protein homolog precursor (Novb) (Nephroblastoma overexpressed
DE gene protein homolog).
GN NOV.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94336229; PubMed=7520150;
RA Martinie C., Huif V., Joubert I., Badzloch M., Saunders G.F.,
RA Strong L.C., Perbal B.;
RT "Structural analysis of the human nov proto-oncogene and expression
RT in Wilms tumor."
RL Oncogene 9:2729-2732(1994).
[2]
RP SEQUENCE FROM N.A.
RA Jiang D., Gou D., Li W.;

```

```

RT "Cloning, sequencing and expression of human nov gene."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.T., Wang J., Hsien F.,
RA Diatchenko L., Marusla K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.D., Hultk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schnerich A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP INTERACTION WITH FBLN1.
RX PubMed=9927660;
RA Perbal B., Martinie C., Salason R., Werner M., He B., Roizman B.;
RT "The C-terminal domain of the regulatory protein NOVb is sufficient to
RT promote interaction with fibulin 1C: a clue for a role of NOVb in
RT cell-adhesion signaling."
RL Proc. Natl. Acad. Sci. U.S.A. 96:869-874(1999).
CC - FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
CC GROWTH REGULATION (BY SIMILARITY).
CC - SUBUNIT: Interacts with FBLN1.
CC - SUBCELLULAR LOCATION: Secreted.
CC - TISSUE SPECIFICITY: INCREASED EXPRESSION IN WILMS TUMOR OF THE
CC STROMAL TYPE.
CC - SIMILARITY: Contains 1 IGFBP domain.
CC - SIMILARITY: Contains 1 VMFC domain.
CC - SIMILARITY: Contains 1 TSP type-1 domain.
CC - SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X78351; CAAS5146.1; -
DR EMBL: X78352; CAAS5146.1; JOINED.
DR EMBL: X78353; CAAS5146.1; JOINED.
DR EMBL: X78354; CAAS5146.1; JOINED.
DR EMBL: X96584; CAAS5403.1; -
DR EMBL: AY082381; AA192490.1; -
DR EMBL: BC015028; AA192490.1; -
DR PIR: I38069; I38069.
DR Genew: HGNC:7885; NOV.
DR MTM: 164958;
DR InterPro: IPR006208; Cys_knot.
DR InterPro: IPR006207; Cys_knot_C.
DR InterPro: IPR000867; Insi_gro_fac_pr.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR001007; VMFC.
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; TSP_1; 1.
DR Pfam: PF00093; vmc; 1.
DR SMART: SM00041; Ctr; 1.

```



```

OY 181 ---KELGFDAEVELLRNNELIANGKGRSLKRLPVFGMEPRILYNPLGOKCIYOTTSMS 237
    |||||
DB 175 RPEATLGIDASDTSR-----CIAOTTEMS 199

OY 238 QCSKTGCTGISTRTVNDNEPCRLVETRICEVKRC-GOPYVSSLKGRKCKSKTKSPRPV 296
    |||||
DB 200 ACSKTCGMSVSSRPTNRNARCEMOKQIRLCVNRSEEPGMEHVEKKKCVKRTTKPI 259

OY 237 RFTYAGCLSVKRYKRYKCGSCVDGRCCTPQLTRIVYKMFRCEDGETFSKNVMIOSCKCN 356
    |||||
DB 260 HFHYKNTSTVQPYKPKFCGCSGDCRCPSTHSTKTHVFEVCPQKRIYKVPVAVISTVCVCH 319

OY 357 YNCP-----HANEAPR 368
    |||||
DB 320 YNCPQDSLSLOYENARFP 337

RESULT 14
NOV_MOUSE STANDARD: PRT: 354 AA.
AC 064293;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE NOV protein homolog precursor (NOVH) (Nephroblastoma overexpressed
  gene protein homolog).
GN NOV.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv, and ICR; TISSUE=Brain;
RX MEDLINE=97131523; PubMed=8975721;
RA Smith M.R., Natarajan D., Taylor L.B., Choi C.P., Martinerie C.,
RA Perbal B., Schofield P.N., Boulter C.A.;
RT "Genomic structure and chromosomal mapping of the mouse nov gene.";
RT Genomics 38:425-428(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=96204003; PubMed=8622864;
RA Martinerie C., Chevalier G., Kauscher F.J. III, Perbal B.;
RT "Regulation of nov by Wt1: a potential role for nov in
  nephrogenesis.";
RT Oncogene 12:1479-1492(1996).
RL -1- FUNCTION: IMMEDIATE-EARLY PROTEIN LIKELY TO PLAY A ROLE IN CELL
CC -1- GROWTH REGULATION (BY SIMILARITY).
CC -1- SUBUNIT: Interacts with FBLN1 (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Contains 1 IGFBP domain.
CC -1- SIMILARITY: Contains 1 WFCC domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X97863; CAAG6457.1; -
DR EMBL: Y09257; CAAG70454.1; -
DR EMBL: X96585; CAAG5404.1; -
DR MGI: 109185; NOV.
DR InterPro: IPR006208; Cys_knot;
DR InterPro: IPR006207; Cys_knot_C;
DR InterPro: IPR000867; Ins1_gro_fac_pr;
DR InterPro: IPR000884; TSP1.

```

```

DR InterPro: IPR001007; WVF_C;
DR Pfam: PF00007; Cys_knot; 1.
DR Pfam: PF00219; IGFBP; 1.
DR Pfam: PF00090; TSP_1; 1.
DR Pfam: PF00093; WVC; 1.
DR SMART: SM00041; CT; 1.
DR SMART: SM00121; IB; 1.
DR SMART: SM00209; TSP1; 1.
DR SMART: SM00214; WVC; 1.
DR PROSITE: PS01185; CTCK_1; 1.
DR PROSITE: PS01225; CTCK_2; 1.
DR PROSITE: PS00222; IGF_BINDING; 1.
DR PROSITE: PS00092; TSP1; 1.
DR PROSITE: PS01208; WVF_C_1; 1.
DR PROSITE: PS0184; WVF_C_2; 1.
DR Growth factor; Signal.
KW SIGNAL 1
FT CHAIN 22 354 POTENTIAL.. NOV PROTEIN HOMOLOG.
FT DOMAIN 29 101 IGFBP.
FT DOMAIN 102 168 WVF_C.
FT DOMAIN 202 247 TSP TYPE-1.
FT DOMAIN 261 335 CTCK.
FT DISULFID 261 298 BY SIMILARITY.
FT DISULFID 278 312 BY SIMILARITY.
FT DISULFID 289 328 BY SIMILARITY.
FT DISULFID 292 330 BY SIMILARITY.
FT DISULFID 297 334 BY SIMILARITY.
FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 354 AA; 38928 MW; 08CE8CFC67829DE CRC64;

Query Match 38.8%; Score 822; DB 1; Length 354;
Best Local Similarity 44.5%; Pred. No. 9.5e-53;
Matches 162; Conservative 46; Mismatches 120; Indels 36; Gaps 9;

OY 7 RALALVYLTLH-LTRFLAS--TCPAAC--HCPLEAPKCAPGVGYVRDGCCKVCAPQLNE 62
    |||||
DB 8 RCLCLGFLFLHLSVGSASLKPSPKPCISPTCAPGRSVYLDGSCCPYCARQNGE 67

OY 63 DCSKTQPCDHTKGLCECNFGASSTALGICRAOSEGRPCENRSRIYONGSEFQPCQHCT 122
    |||||
DB 68 SCSEMRPCDQSSGFLYCDRSADPNNGTICMV--PEGDNCVPGVIRNKEKEFEPNQYRCT 126

OY 123 CIDGAVGICPLCPQELSPINLGCPRRLVYKTYTGCCCEWVDEDSIKRPMEDQGLKE 182
    |||||
DB 127 CRDQIGCLPRLCQDLVLLPDPCCAPRAVAPGECCEKWTGSDS-----CGTGTLG-G 180

OY 183 LGFPAEVELLRNNELIANGKGRSLKRLPVFGMEPRILYNPLGOKCIYOTTSMSQSKT 242
    |||||
DB 181 LALRAYRPEAR-----VGVEVSDSSI-----NCIEOTTMSAGSKS 216

OY 243 CGTGISTRTVNDNEPCRLVETRICEVKRCG-PYSSSLKGRKCKSKTKSPRPVFTYA 301
    |||||
DB 217 CGMGVSTRTVNRNRCCEWVKQTRICIVAPCDQEPREYTDKKGKCLFRKSLKAIHLQFE 276

OY 302 GCLSVKRYKRYKCGSCVDGRCCTPQLTRIVYKMFRCEDGETFSKNVMIOSCKKNYPH 361
    |||||
DB 277 NCTSLTYKPRKPCGVCSDGRCCTPHNTKTIOVEQCLPGLGELIKKEVAVIGTCYCSNCPQ 336

OY 362 ANEA 365
    |||||
DB 337 NNEA 340

RESULT 15
NOV_RAT
ID NOV_RAT STANDARD: PRT: 351 AA.
AC 090205;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE NOV protein homolog precursor (NOVH) (Nephroblastoma overexpressed
  gene protein homolog).

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 5, 2003, 14:02:31 ; Search time 33.0868 Seconds
(without alignments)
2971.515 Million cell updates/sec

Title: US-09-495-448a-4
Perfect score: 2116
Sequence: 1 MSSRIARALALVYTLHLTR.....ANEAPFPYRLPNDIKFRD 381

Scoring table:
BIOSUM62
Gapop 10.0 , Gapext 0.5*

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhc:*
- 8: sp.organelle:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp.podent:*
- 12: sp.virus:*
- 13: sp.vertebrate:*
- 14: sp.unclassified:*
- 15: sp.virus:*
- 16: sp.bacteriap:*
- 17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1928	91.1	379	11	Q9WTM9
2	1663.5	78.6	375	13	Q98TX5
3	1587	75.0	334	4	Q9UID7
4	938	44.3	343	13	Q42607
5	930	44.0	344	13	Q98T08
6	912.5	43.1	347	13	Q9PT80
7	878	41.5	349	6	Q97765
8	822	38.8	354	11	Q8CA67
9	767	36.2	367	11	Q95388
10	758	35.8	367	11	Q99P00
11	758	35.8	367	11	Q54775
12	626.5	29.6	331	4	Q95958
13	613	29.0	354	4	Q95389
14	514.5	24.3	251	11	Q8CIC8
15	495.5	23.4	280	4	Q9HCS3
16	485.5	22.9	176	13	Q9PS56

17	373.5	17.7	128	11	Q9R2C0	Q9R2C0
18	341	16.1	113	11	Q9Z164	Q9Z164
19	337	15.9	100	4	Q9UD16	Q9UD16
20	336	15.9	119	11	Q91V29	Q91V29
21	336	15.9	119	11	Q920W6	Q920W6
22	302	14.3	374	5	Q9VVK0	Q9VVK0
23	294	13.9	470	5	Q9VVK3	Q9VVK3
24	293	13.8	230	4	Q8WYK7	Q8WYK7
25	195	9.2	77	4	Q9UDE4	Q9UDE4
26	187	8.8	2327	13	Q91BG7	Q91BG7
27	176	8.3	70	13	Q9DF21	Q9DF21
28	174	8.2	1045	5	Q8T3A6	Q8T3A6
29	174	8.2	1070	5	Q8T3A7	Q8T3A7
30	174	8.2	1111	5	Q9XMD6	Q9XMD6
31	170.5	8.1	1036	4	Q9NZV1	Q9NZV1
32	169	8.0	1664	5	Q9TV02	Q9TV02
33	169	8.0	1704	5	Q94446	Q94446
34	168.5	8.0	1028	11	Q9JIL0	Q9JIL0
35	166.5	7.9	1048	13	Q8AW55	Q8AW55
36	164	7.8	58	6	Q97574	Q97574
37	162.5	7.7	4998	11	Q8CG65	Q8CG65
38	161.5	7.6	2104	5	Q21281	Q21281
39	161.5	7.6	2104	5	Q964N4	Q964N4
40	161	7.6	360	5	Q8MN64	Q8MN64
41	159.5	7.5	406	11	Q92513	Q92513
42	159.5	7.5	426	11	Q8VZA6	Q8VZA6
43	159	7.5	5146	6	Q8SPW4	Q8SPW4
44	158.5	7.5	792	13	Q90Z43	Q90Z43
45	158.5	7.5	1095	13	Q90XG4	Q90XG4

ALIGNMENTS

RESULT 1

ID	Q9WTM9	PRELIMINARY:	PRT:	379 AA.
AC	Q9WTM9			
DT	01-NOV-1999 (TREMBLrel. 12, Created)			
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	CYR61 precursor.			
GN	CYR61.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_Taxid=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-Item; TISSUE-Aorta;			
RA	Unoki H., Yonekura H., Furukawa K., Yamamoto H.;			
RT	"Rat CyR61 mRNA."			
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AB015877; BAA78339.1; -			
DR	InterPro: IPR006208; Cys_knot.			
DR	InterPro: IPR006207; Cys_knot_C.			
DR	InterPro: IPR000867; Insl_gro_fac_pr.			
DR	InterPro: IPR000884; TSPI.			
DR	InterPro: IPR001007; VWF_C.			
DR	Pfam: PF00007; Cys_knot; 1.			
DR	Pfam: PF00219; IGFBP; 1.			
DR	Pfam: PF00090; tsp_1; 1.			
DR	Pfam: PF00093; vwc; 1.			
DR	SMART: SM00041; CT; 1.			
DR	SMART: SM00209; TSPI; 1.			
DR	SMART: SM00212; IB; 1.			
DR	SMART: SM00214; VMC; 1.			
DR	PROSITE: PS01185; CTCK_1; 1.			
DR	PROSITE: PS01225; CTCK_2; 1.			
DR	PROSITE: PS00222; IGF_BINDING; 1.			
DR	PROSITE: PS50092; TSPI; 1.			
DR	PROSITE: PS01208; VWF_C; 1.			
KW	Signal.			

FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 379 CYR61.
 SQ SEQUENCE 379 AA; 41728 MW; D2A8AFD7B84762B CRC64;
 Query Match 91.18; Score 1928; DB 11; Length 379;
 Best Local Similarity 90.6%; Pred. No. 2,8e-179;
 Matches 347; Conservative 11; Mismatches 19; Indels 6; Gaps 2;

QY 1 MSSRIKALAVYTLHLTRIALSTCPACCHCPLKAPKAPGVGLVRDGGCCAKOL 60
 1 MSSSTIKTLAVNTLHLTRIALSTCPACCHCPLKAPKAPGVGLVRDGGCCAKOL 60
 DB 1 MSSSTIKTLAVNTLHLTRIALSTCPACCHCPLKAPKAPGVGLVRDGGCCAKOL 60
 QY 61 NEDSCSTQPCDHFKGLKCNFAGSSTALKGICRAOSEGRPCENSRITQNGSEFQPCNQ 120
 61 NEDSCSTQPCDHFKGLKCNFAGSSTALKGICRAOSEGRPCENSRITQNGSEFQPCNQ 120
 DB 61 NEDSCSTQPCDHFKGLKCNFAGSSTALKGICRAOSEGRPCENSRITQNGSEFQPCNQ 120
 QY 121 CTCIDAVGCIPLCPDELSPNLGCPNRLVKTGOCCEWVDESDIKDLDODDL 180
 121 CTCIDAVGCIPLCPDELSPNLGCPNRLVKTGOCCEWVDESDIKDLDODDL 180
 DB 121 CTCIDAVGCIPLCPDELSPNLGCPNRLVKTGOCCEWVDESDIKDLDODDL 180
 QY 181 KELGPDASEVELTRNNELIANGKRSRLKPLVPGMPPRLYNPL--OGOKCIYQTSMSQ 238
 179 --LGFDASEVELTRNNELIATGKSSLKRLPVGTEPRVLYNPLAHGOKCIYQTSMSQ 236
 DB 179 --LGFDASEVELTRNNELIATGKSSLKRLPVGTEPRVLYNPLAHGOKCIYQTSMSQ 236
 QY 239 CSKTCGTGISTRTNDNPECRILVKEPRICEVRPCGPVYSSLKGGKSKTKKSPPEVR 298
 237 CSKSCGTGISTRTNDNPECRILVKEPRICEVRPCGPVYSSLKGGKSKTKKSPPEVR 296
 DB 237 CSKSCGTGISTRTNDNPECRILVKEPRICEVRPCGPVYSSLKGGKSKTKKSPPEVR 296
 QY 299 TYAGCISVKKRYKYGSCVDGRCCTPOLTRVTKMFRCEDETFESKNVMIOSCKCN 358
 297 TYAGCISVKKRYKYGSCVDGRCCTPOLTRVTKMFRCEDETFESKNVMIOSCKCN 356
 DB 297 TYAGCISVKKRYKYGSCVDGRCCTPOLTRVTKMFRCEDETFESKNVMIOSCKCN 356
 QY 359 CPHANEAPFYRLFNDIHKFRD 381
 357 CPHANEAPFYRLFNDIHKFRD 379
 DB 357 CPHANEAPFYRLFNDIHKFRD 379

RESULT 2
 Q98TX5 PRELIMINARY; PRT; 375 AA.
 ID Q98TX5
 AC Q98TX5
 DT 01-JUN-2001 (TEMBLrel. 17, Created)
 DT 01-JUN-2001 (TEMBLrel. 17, last sequence update)
 DT 01-MAR-2003 (TEMBLrel. 23, last annotation update)
 DE secreted cysteine-rich protein cyr61.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodidae; Xenopus.
 NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Latick B.V., Benett B., Smith J.C.;
 RT Characterization of Xenopus cyr61.
 RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF20592; AAK00947.1;
 DR EMBL; AF20592; AAK00947.1;
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot.C.
 DR InterPro: IPR000867; Insl_gro_fac-pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF000219; IGFBP; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR Pfam: PF00093; vwc; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF_BINDING; 1.
 DR PROSITE; PS00922; TSP1; 1.

DR PROSITE; PS01208; VWF_C; 1.
 SQ SEQUENCE 375 AA; 41460 MW; 78075CA7B380304E CRC64;
 Query Match 78.6%; Score 1663.5; DB 13; Length 375;
 Best Local Similarity 79.3%; Pred. No. 1.5e-153;
 Matches 295; Conservative 29; Mismatches 41; Indels 7; Gaps 5;

QY 11 LVYTLH-LTRLSTCPACCHCPLKAPKAPGVGLVRDGGCCAKOLNEDSKTOP 69
 11 LVYTLH-LTRLSTCPACCHCPLKAPKAPGVGLVRDGGCCAKOLNEDSKTOP 69
 DB 11 LVYTLH-LTRLSTCPACCHCPLKAPKAPGVGLVRDGGCCAKOLNEDSKTOP 69
 QY 10 LALLSGFLDLAVSSCPAVCCPVEPPKAPGVGLVDGCGCKCAQOLNEDSKTHP 69
 10 LALLSGFLDLAVSSCPAVCCPVEPPKAPGVGLVDGCGCKCAQOLNEDSKTHP 69
 DB 10 LALLSGFLDLAVSSCPAVCCPVEPPKAPGVGLVDGCGCKCAQOLNEDSKTHP 69
 QY 70 CDHTKLECNFAGSSTALKGICRAOSEGRPCENSRITQNGSEFQPCNQCTCIDGAVG 129
 70 CDHTKLECNFAGSSTALKGICRAOSEGRPCENSRITQNGSEFQPCNQCTCIDGAVG 129
 DB 70 CDHTKLECNFAGSSTALKGICRAOSEGRPCENSRITQNGSEFQPCNQCTCIDGAVG 129
 QY 130 CIPLCFQELSLPLGCPNRLVKTGOCCEWVDESDIKDLDODDL 189
 130 CIPLCFQELSLPLGCPNRLVKTGOCCEWVDESDIKDLDODDL 187
 DB 130 CIPLCFQELSLPLGCPNRLVKTGOCCEWVDESDIKDLDODDL 187
 QY 190 VELTRNNELIANGKRSRLKPLVPGMPPRLYNPLDLOGKCIYQTSMSQCTGTGIST 249
 188 VELTRNNELIANGKRSRLKPLVPGMPPRLYNPLDLOGKCIYQTSMSQCTGTGIST 244
 DB 188 VELTRNNELIANGKRSRLKPLVPGMPPRLYNPLDLOGKCIYQTSMSQCTGTGIST 244
 QY 250 RVTNDNPECRILVKEPRICEVRPCGPVYSSLKGGKSKTKKSPPEVRRTYAGLSVYKT 309
 245 RVTNDNPECRILVKEPRICEVRPCGPVYSSLKGGKSKTKKSPPEVRRTYAGLSVYKT 304
 DB 245 RVTNDNPECRILVKEPRICEVRPCGPVYSSLKGGKSKTKKSPPEVRRTYAGLSVYKT 304
 QY 310 RPKYGGSCVDGRCCTPOLTRVTKMFRCEDETFESKNVMIOSCKCNYPHANEAPF 369
 305 RPKYGGSCVDGRCCTPOLTRVTKMFRCEDETFESKNVMIOSCKCNYPHANEAPF 363
 DB 305 RPKYGGSCVDGRCCTPOLTRVTKMFRCEDETFESKNVMIOSCKCNYPHANEAPF 363
 QY 370 YRLFNDIHKFRD 381
 364 YRLFNDIHKFRD 375
 DB 364 YRLFNDIHKFRD 375

RESULT 3
 Q9UID7 PRELIMINARY; PRT; 334 AA.
 ID Q9UID7
 AC Q9UID7
 DT 01-MAY-2000 (TEMBLrel. 13, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, last sequence update)
 DT 01-MAR-2003 (TEMBLrel. 23, last annotation update)
 DE cyr61 protein.
 GN cyr61.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 1-107 FROM N.A.
 RA Anding B., Long Y.;
 RT Tissue-Kidney;
 RT "Cloning of a new gene down-regulated in the small-cell tumor
 embryonal-rhabdomyosarcoma (RMS)."
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF003114; AAF21597.1;
 DR EMBL; AF003114; AAF21597.1;
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot.C.
 DR InterPro: IPR000867; Insl_gro_fac-pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF000219; IGFBP; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR Pfam: PF00093; vwc; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS01208; VWF_C; 1.
 DR PROSITE; PS01208; VWF_C; 1.
 SQ SEQUENCE 334 AA; 37246 MW; 9188987A7352E948 CRC64;

Query Match 75.0%; Score 1587; DB 4; Length 334;
 Best Local Similarity 85.8%; Pred. No. 3.8e-146;
 Matches 295; Conservative 5; Mismatches 22; Indels 22; Gaps 4;

49 GCGCCKVCACQALNEDCSKTQPODHTKGLGECNFGASSTALGICRAOSEGRCEVNSRTYQ 108
 2 GCGTHPNLCIHGHTASPTSYKHHKLECNFGASSTALGICRAOSEGRCEVNSRTYQ 61
 109 NGESQPNQCQHCCTCIDAVAGCIPLCPELSLPLNIGCNPRYLKVTGGCCCEWVCDESI 168
 62 NGESQPNCKHQCTCIDAVAGCIPLCPELSLPLNIGCNPRYLKVTGGCCCEWVCDESI 121
 169 KDPMDODGLGKELGEPFASVEVELTRNNELIYAVGKSLKRLPYVGMPEPRLIYNLOGOK 228
 122 KDPMDODGLGKELGEPFASVEVELTRNNELIYAVGKSLKRLPYVGMPEPRLIYNLOGOK 181
 229 CIVQTTSMQSCSKTCGTGISTRTVNDNEPCRLVETRICEVNPGQPYVSSLKRGKCSK 288
 182 CIVQTTSMQSCSKTCGTGISTRTVNDNEPCRLVETRICEVNPGQPYVSSLKRGKCSK 241
 289 TKKSEPEVRYTYAGCLSKTKRKYKCGSCVDRCTPOLITRTVMKRFCEDE-----GE 341
 242 TKKSEPEVRYTYAGCLSKTKRKYKCGSCVDRCTPOLITRTVMKRFCEDE-----GE 341
 342 TFSKN-----VMMIOSCKNCYNCPHANEAPFFYRLFNDFIHKFRD 381
 296 IFQERHDDPVL-----KCNYNCPHANEAPFFYRLFNDFIHKFRD 334

Db

RESULT 4

042607 PRELIMINARY; PRT: 343 AA.

AC 042607; PRELIMINARY; PRT: 343 AA.

DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Connective tissue growth factor XCTGF.
 OS Xenopus laevis (African clawed frog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ying Z., King M.L.;
 RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U43524; AAB67639.1; -
 DR EMBL; U43523; AAB67638.1; -
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Insl_gro_fac.pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00219; IGFBP; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR Pfam: PF00093; vwc; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS0092; TSP1; 1.
 DR PROSITE; PS01208; WFC; 1.
 SO SEQUENCE 343 AA; 37966 MW; 93F221C5DB565A81 CRC64;

Query Match 44.3%; Score 938; DB 13; Length 343;
 Best Local Similarity 47.2%; Pred. No. 5.8e-83;
 Matches 168; Conservative 51; Mismatches 95; Indels 42; Gaps 6;

26 CPACHCPLEAPKCAPGVGLVDRGCGCCKVCACQALNEDCSKTQPCDHTKGLGECNFGASST 85

Db

24 CNGEQCNKPYVQDGPVARMVQDGGCCCKVSCQKGLGELCTERDVCDPHKGLFCDRGSRVN 83
 86 ALKICRAOSEGRCEVNSRTYQNGESQPNQCQHCCTCIDAVAGCIPLCPELSLPLNIGC 145
 84 RKIGVCTAR-EGAPCEVFGVYRSGESFQSSCKQCTCIDAVGCVPLCSMDIRLPSPEC 142
 146 PNPRLVKTGGCCCEWVCDESIDPMDODGLGKELGEPFASVEVELTRNNELIYAVGKGR 205
 143 PPNRVRKLPKGCCEWVCDESIDPMDODGLGKELGEPFASVEVELTRNNELIYAVGKGR 172
 206 SLKRLPYVGMPEPRLIYNLOGOKCIVQTTSMQSCSKTCGTGISTRTVNDNEPCRLV 262
 173 -----LPARME--TYGDPPLIRANCLVQTTSMQSCSKTCGTGISTRTVNDNEPCRLV 226
 263 ETRICEVNPQPYVSSLKRGKCSKTKKSEPEVRYTYAGCLSKTKRKYKCGSCVDRGRC 322
 227 QSRICMVRPCADIEENIKKCKIRPKIRKPVKFEFGTQSVKTYRAKCGCTDGRG 286
 323 CTQPLTRTVKMFCEDETFKSNVMMIOSCKNCYNCPHANEAPFFYRLFNDFI 376
 287 CTPHRTATLPVEFKCPDGEVKKMMFTKTCACHNCPGDDNIFEAMTYRKATGDM 342

Db

RESULT 5

098T08 PRELIMINARY; PRT: 344 AA.

AC 098T08; PRELIMINARY; PRT: 344 AA.

DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Connective tissue growth factor precursor (Connective tissue growth factor/hypertrophic chondrocyte-specific protein 24).
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gyi D.;
 RL Submitted (Jan-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Mukunda Y., Kubota S., Takigawa M.;
 RL Submitted (Dec-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ298335; CAC33438.1; -
 DR EMBL; AF463517; AAL68834.1; -
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot_C.
 DR InterPro: IPR000867; Insl_gro_fac.pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00219; IGFBP; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR Pfam: PF00093; vwc; 1.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00121; IB; 1.
 DR SMART; SM00209; TSP1; 1.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00222; IGF_BINDING; 1.
 DR PROSITE; PS0092; TSP1; 1.
 DR PROSITE; PS01208; WFC; 1.
 KW SIGNAL.
 FT SIGNAL 1 21
 FT CHAIN 22 344
 SO SEQUENCE 344 AA; 37499 MW; 69E639AF69BFLD00 CRC64;

Query Match 44.0%; Score 930; DB 13; Length 344;
 Best Local Similarity 45.3%; Pred. No. 3.5e-82;

Matches	169:	Conservative	59:	Mismatches	105:	Indels	40:	Gaps	7:
QY	8	ALALVTLLHLTRALSTCPACHCPL-EAPKCAPGVGLVRDGGCCCKVCYAKOLNEDSCS	66						
Db	7	AVALLLALLIG-PEVAGGECGCGCCGSGGPGSPCAGVSLVLDGCGCCRVCAKOLGELCTE	65						
QY	67	TOPDHTGLGCLNGCNGASSTALKIGICRQASRSEGRPCENSRITONESHSPQPNQHOCTCIDG	126						
Db	66	RDPCDHRHGLCLCDRSPSNRRIGVCTAN-DAAPCVFSQMYVRSGSFSSCKYQCTCIDG	124						
QY	127	AVGCIPLCPQELSLPNLGCNPRLVYKYGCCCEEWVCCDEDSIKDPMEDQDGLLKEIGFD	186						
Db	125	AVGCVPLCSMVRRLRSPDPCFPFRKKLPGKCCCEWVCCDE-----	163						
QY	187	ASEVELFRNNELIAYGKRSIKRL-PVFGMEPRILYNLQOKCIVQTTSMSSQSKTGT	245						
Db	164	-----AKBQTAAGPALAAYRLIEDFTGDPPTMM-----RANCLVQTTEWSAKSTGGM	210						
QY	246	GISTRTVTNDNEPCRLVYKTRICEVRRPGQAPVYSSLKKGKSKTKKSPRPFTIYAGLS	305						
Db	211	GISTRTVNDNAFCRIEKSRLCMVRPCADLEENIKKKKICIRPPKISKPIKFLSGCTS	270						
QY	306	VKKTRPKYKCGVDRGCTPOLFTRTVKRPFCEGDEFFSKVMMMIOSCKVYNCPHANEA	365						
Db	271	VKTRAFKCGCTGDRCTPHRTATLIVFEFKCPDGEIKRKMMPKITCACHYNGPDNDI	330						
QY	366	AFPEY-RLFNDI 376							
Db	331	FESLYRRKMYGDM 343							
RESULT 6									
ID	Q9PT80	PRELIMINARY;	PRT;	347	AA.				
AC	Q9PT80;								
DT	01-MAY-2000 (TREMBLrel. 13, Created)								
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)								
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)								
DE	Connective tissue growth factor.								
GN	CTGF.								
OS	Notophthalmus viridescens (Eastern newt) (Triturus viridescens).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;								
OC	Notophthalmus.								
OX	NCBI_TaxID=8316;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Forelimb blastema;								
RA	Gates P.B.;								
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.								
RN	[2]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Forelimb blastema;								
RA	Cash D.E.; Gates P.B.; Imokawa Y.; Brookes J.P.;								
RT	"Identification of newt connective tissue growth factor as a target of								
RT	retinoid regulation in limb blastemal cells.";								
RL	Gene 222:119-124(1998).								
DR	EMBL; AJ271167; CAB65965.1; -								
DR	InterPro: IPR006208; Cys_knot.								
DR	InterPro: IPR006207; Cys_knot_C.								
DR	InterPro: IPR000867; Insl_gro_fac_pr.								
DR	InterPro: IPR000884; TSP1.								
DR	InterPro: IPR001007; WWF_C.								
DR	Pfam: PF000007; Cys_knot.1.								
DR	Pfam: PF00219; IGFBP.1.								
DR	Pfam: PF00090; tsp_1.1.								
DR	Pfam: PF00093; wvc_1.								
DR	SMART: SM00041; CT.1.								
DR	SMART: SM00121; IB.1.								
DR	SMART: SM00209; TSP1.1.								
DR	SMART: SM00214; WVC.1.								
DR	PROSITE; PS01185; CTCK_1.1.								

Query Match	43.1%	Score 912.5	DB 13	Length 347
Best Local Similarity	45.1%	Pred. No. 1.8e-80		
Matches 167	Conservative 54	Mismatches 112	Indels 37	Gaps 6
OY	9	LALVYTLHLRLALSTCPACHCPLLEAPKCAPGVGLVRDGGCGCCVCAPKOLNEDDCSKTQ	68	
DB	12	LLIALLSWSCA-QDCSGECRCRNKRPCEPAGISLVMDGCGCCVCAPKOLGELCTEKD	70	
OY	69	PCDHATGECNFGASSTALKGICRAQSGRPCEVRSRIYQNGESPQNCOCCTCIDAV	128	
DB	71	VCDPHRGLFCDFSGSVNKKKIGVCTAK-DGAPCVFGVYRSGESQSSCKYGCCTLDGCV	129	
OY	129	GCIPCLCPDELSPNIGCNPRLVAVYTGCCCEWVCDDEDSIKDPMEDQGLGKELGFDAS	188	
DB	130	GCVPPLCGVDVRLPSPDCFPFRVVKOLGKCEWVCDQK-----	168	
OY	189	EVELTRNNELLAVGSGRLKRLPVGSGPRILYLNPLQOQKCIYOTTSMSCKTGTGTIS	248	
DB	169	--EOTRGPAIAV-----YRDETYGPD-----SSLRANCLVOTTESASCKTGMGIS	216	
OY	249	TRVINDNECRCLVKEATRICEVAPCGQPYSSSLKNGKCKSKTKKSPPEVRYAGLSYK	308	
DB	217	TRVINDNEMCKLERKOSRLCMVRPCADLEENIKKKKICIRPKISKIPKFLSGTSVKT	276	
OY	309	YRPKYCGSCVDRCTPOLITRYTKMRFCEDGEFTSKVNMVMIQSCKCYNCPHNEAAP	368	
DB	277	YRAFECGVCYCTDGRCTCPHRTATLPLPEFKPCDPGEVVKMKMFITKACAHYNGDNDIFES	336	
OY	369	FY--RLFNDI 376		
DB	337	MYRKMTGDM 346		

RESULT 7

097765 PRELIMINARY: PRT: 349 AA.

ID	097765	PRELIMINARY:	PRT:	349 AA.
AC	01-MAY-1999	(TREMBLrel. 10, Created)		
DT	01-MAY-1999	(TREMBLrel. 10, Last sequence update)		
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)		
DE	Connective tissue growth factor.			
DE	CTGF.			
OS	Sus scrofa (pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=uterus;			
RA	Harding P.A., Brigstock D.R.;			
RT	"Cloning and sequencing of a porcine connective tissue growth factor			
RT	(CTGF) cDNA."			
RL	Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.			
DR	EMBL, U70060, AAD00174.1, -			
DR	InterPro: IPR006208; Cys_Knot.			
DR	InterPro: IPR006207; Cys_Knot_C.			
DR	InterPro: IPR000867; Insl_gro_fac_pr.			
DR	InterPro: IPR000884; TSP1.			
DR	InterPro: IPR001007; VWF_C.			
DR	Pfam: PF00007; Cys_Knot; 1.			
DR	Pfam: PF00219; IGFBP; 1.			
DR	Pfam: PF00090; tsp_1; 1.			
DR	Pfam: PF00093; VWC; 1.			
DR	SMART: SM00041; CT; 1.			
DR	SMART: SM00121; IB; 1.			
DR	SMART: SM00209; TSP1; 1.			
DR	SMART: SM00214; VWC; 1.			

DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS01208; VMEC; 1.
SQ SEQUENCE 367 AA; 40331 MW; 9F29CA94D69C0502 CRC64;

Query Match 36.2%; Score 767; DB 4; Length 367;
Best Local Similarity 38.0%; Pred. No. 2.8e-66;
Matches 148; Conservative 62; Mismatches 125; Indels 54; Gaps 9;

QY 2 SSRIRALALVLTLLHTRIALST-----CPACHCPLEAPKCAPGVLRDGGCGCK 54
DB 18 STVALATSPAPPTMTFTFAPLEITREPECKMPCCECPSPRCPGLVSLITDGECK 77
QY 55 VCAKQINEDCSKTOPCDHTKGLCNF--GASSTALKGICRAGSEGRPCENSRITONGSEF 113
DB 78 MCAQOIGDNCCTEALIDDPHRLGYCDYSGDRPRYAIGVC--AQVVGVCVLDGVRYYNGQSF 136
QY 114 QPACQHCCTCIDAVAGCIPCLPQELSLPNLGCNPNRLVKTGCGCEWYCDDESDIKDPM 173
DB 137 QPCKKNCCTCIDAVAGCIPCLPQELSLPNLGCNPNRLVKTGCGCEWYCDDESDIKDPM 194
QY 174 DODGLGKELGEPDA--SEVELTRNNELIAGKGRSLKRLPVGMEPRILYNPLQOGKCIYO 232
DB 195 TAPRDGT---AFDAVEVEAMHNR-----CIAY 219
QY 233 TTSWQSCSKTCGTGISTRTVNDNPECRIVKTRICEVRCGOPYSSLLKKGKSKTKS 292
DB 220 TTSWQSCSKTCGTGISTRTVNDNPECRIVKTRICEVRCGOPYSSLLKKGKSKTKS 279
QY 293 PEFVRYTVAAGLSVKKRYRKYCGSCVDGRCCTPOLRTVKKRFRCEDETFSKNMMIOS 352
DB 280 EASMTTLGACISTRTVNDNPECRIVKTRICEVRCGOPYSSLLKKGKSKTKS 339
QY 353 CKNYCNPHANEAAPFYRLFNDIHKFRD 381
DB 340 CECNLSCRPNND-----IFADLESYPD 361

RESULT 10

Q99PPO PRELIMINARY; PRT; 367 AA.
AC Q99PPO;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE ELM1.
GN ELM1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20487548; PubMed=11031104;
RA Sleeman M.A., Murlison J.G., Strachan L., Kumbale K.D., Glenn M.P.,
RA McGrath A., Bickertstaff P., Grierson A., Havukala I., Tan P.,
RA Watson J.D.;
RT "Gene expression in rat dermal papilla cells: analysis of 2529 ESTs";
RL Genomics 63;214-224(2000).
DR EMBL; AF228045; AAK00729.1; -
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot.C.
DR InterPro; IPR000867; Insl_gro_fac.Pr.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR001007; VME_C.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF00219; IGFBP; 1.
DR Pfam; PF00090; tsp; 1; 1.
DR Pfam; PF00093; VME; 1.
DR SMART; SM00041; Cyt; 1.
DR SMART; SM00121; IB; 1.
DR SMART; SM00209; TSP1; 1.
DR SMART; SM00214; VME; 1.
DR PROSITE; PS01185; CTCK_1; 1.

DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS01208; VMEC; 1.
SQ SEQUENCE 367 AA; 40613 MW; 8A4A3AC69D3243D2 CRC64;

Query Match 35.8%; Score 758; DB 11; Length 367;
Best Local Similarity 38.5%; Pred. No. 2.1e-65;
Matches 149; Conservative 52; Mismatches 130; Indels 56; Gaps 8;

QY 5 IARALALVLTLLHTRIALST-----CPACHCPLEAPKCAPGVLRDGGCGCKKCA 57
DB 21 LATALSPTPTMTFTFAPLEITREPECKMPCCECPQAPRCPGLVSLITDGECKICA 80
QY 58 KQINEDCSKTOPCDHTKGLCNF--GASSTALKGICRAGSEGRPCENSRITONGSEFQPN 116
DB 81 QOLGDNCCTEALIDDPHRLGYCDYSGDRPRYAIGVC--AQVVGVCVLDGVRYYNGQSF 139
QY 117 COHCCTCIDAVAGCIPCLPQELSLPNLGCNPNRLVKTGCGCEWYCDDESDIKDPM 174
DB 140 CRINCTCIDGTVGCTPLCLSPR---PRLMCRQPRHVRVPGQCEWYCDDESDIKDPM 193
QY 175 ODGLGKELGEPDA--SEVELTRNNELIAGKGRSLKRLPVGMEPRILYNPLQOGKCIYO 234
DB 194 QTALDTRAFASGAVEREN-----CIAYS 221
QY 235 SMSQSCSKTCGTGISTRTVNDNPECRIVKTRICEVRCGOPYSSLLKKGKSKTKSPE 294
DB 222 PMSQSCSKTCGTGISTRTVNDNPECRIVKTRICEVRCGOPYSSLLKKGKSKTKSPE 281
QY 295 PVFVRYTVAAGLSVKKRYRKYCGSCVDGRCCTPOLRTVKKRFRCEDETFSKNMMIOSCK 354
DB 282 ATNFTLAGCISTRTVNDNPECRIVKTRICEVRCGOPYSSLLKKGKSKTKS 341
QY 355 CKNYCNPHANEAAPFYRLFNDIHKFRD 381
DB 342 CKNLSCRPNND-----IFADLESYPD 361

RESULT 11

O54775 PRELIMINARY; PRT; 367 AA.
AC O54775;
DT 01-JUN-1998 (TREMblrel. 06, Created)
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
DE ELM1.
GN ELM1 OR WISPL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HEN;
RX MEDLINE=98119879; PubMed=9449709;
RA Hashimoto Y., Shindo-okada N., Tanai M., Nagamachi Y., Takeuchi K.,
RA Shiroishi T., Toma H., Yokoyama J.;
RT Expression of the ELM1 gene, a novel gene of the CCN (connective
RT tissue growth factor, Cy61/Ce10, and neuroblastoma overexpressed
RT gene) family, suppresses in vivo tumor growth and metastasis of K-1735
RL murine melanoma cells";
RL J. Exp. Med. 187:289-296(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=99061933; PubMed=9843955;
RA Penica D., Swanson T.A., Welsh J.W., Roy M.A., Lawrence D.A., Lee J.,
RA Brush J., Taneyhill L.A., Quire B., Lew M., Watanabe C., Cohen R.L.,
RA Melham M.F., Finley G.G., Quire P., Goddard A.D., Hillan K.J.,
RA Gurney A.L., Botstein D., Levine A.J.;
RT "WSP genes are members of the connective tissue growth factor family
RT that are up-regulated in wt-1-transformed cells and aberrantly
RT expressed in human colon tumors";

RL Proc. Natl. Acad. Sci. U.S.A. 95:14717-14722(1998).
 DR EMBL: AB004873; BAA24949.1; -
 DR EMBL: AF100773; AAC96319.1; -
 DR MGD: MGI:1197008; Wtsp1.
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot.C.
 DR InterPro: IPR000867; Insl_gro_fac_pr.
 DR InterPro: IPR000884; TSP1.
 DR InterPro: IPR001007; WVE_C.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00219; IGFBP; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR Pfam: PF00093; wvc; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00214; IB; 1.
 DR SMART: SM00209; TSP1; 1.
 DR SMART: SM00214; WVC; 1.
 DR PROSITE: PS01185; CTCK_1; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR PROSITE: PS50092; TSP1; 1.
 DR PROSITE: PS01208; WVEC; 1.
 DR SEQUENCE 367 AA; 40702 MW; 387C0569EFAE5E96 CRC64;
 SO
 Query Match 35.8%; Score 758; DB 11; Length 367;
 Best Local Similarity 38.5%; Pred. No. 2,1e-65;
 Matches 149; Conservative 52; Mismatches 130; Indels 56; Gaps 8;
 QY 5 IARLAIVTLHLTRALST-----CPACHCLEAPKCAPGVGLVNDGCGCKVCA 57
 DB 21 IARLSTPTPTMTFTTPALERTTRPERCKWPCPCPSPPRCPLGVLITDGCCECKICA 80
 QY 58 KOLNEDSKTQPCDHTGLECNF--GASSTALKGICRAOSEBRCPEVNSRIYONSESOPN 116
 DB 81 QQLDNCNTEALICPHNGLCYDSGDRPRRAIGVC-AQVVGCVLLGVRTNGESFOPN 139
 QY 117 COHQCTIDGAVGCIPLC--POELSLPNLGGPNRLKVVNGCCCEWNCDEDSIKDPMED 174
 DB 140 CRVCTCIDGVTGCTPLCLSPR---PRMCRQPRNHRVRPQCCCEQWVCDDAR----PR 193
 QY 175 QDGLGKELGDFASEVELTRNNELIAVGKSLRLPFPVGNPRILYNPLQGKCIQVOT 234
 DB 194 QTALLTRAPRAASGAVQREYN-----CIAVTS 221
 QY 235 SMSGCKTGGTGSTRTVNTNPECRLYKEPRICEVRPCGQPVYSLKKGKCKSTKKSPE 294
 DB 222 PWSGCTSTGGLGISTRISNVNARCPQESRLCNLRCDVDIQLHIAAGKCLAVYQPEE 281
 QY 295 PVRETYAGCLSVKKRYKYGSCVDGRCTPOLTRTYKMRPCEDEGETFSKNVMIQSK 354
 DB 282 ATNFTLAGCVSTRYRKYKYGSCVDNRCICPIYKSKTISVDQCPGEGFSGVOYMINACF 341
 QY 355 CNVNCPRANEAAPFRLFENDIHKFRD 381
 DB 342 CNLSCRNPND-----IFADLESYPD 361
 RESULT 12
 095958 PRELIMINARY; PRT; 331 AA.
 AC 095958;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE D014217.3 (Connective tissue growth factor (NOV, GIG) like protein)
 DE (Fragment).
 GN D014217.3 OR LIBC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tubby B.;

RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Mammary gland;
 RA van Golen K.L., Davies S., Wu Z.F., Wang Y.F., Bucana C.D., Root H.,
 RA Chandrasekharappa S., Strawderman M., Ehlier S.P., Metajver S.D.;
 RT "A novel putative IGF-binding, tumor suppressor protein, LIBC, and
 RT Rhc Gpase, are determinants of the inflammatory breast cancer
 RT phenotype";
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z99289; CAB16556.1; -
 DR EMBL: AF143679; AAD31517.1; -
 DR InterPro: IPR006208; Cys_knot.
 DR InterPro: IPR006207; Cys_knot.C.
 DR InterPro: IPR000867; Insl_gro_fac_pr.
 DR InterPro: IPR000884; TSP1.
 DR Pfam: PF00007; Cys_knot; 1.
 DR Pfam: PF00219; IGFBP; 1.
 DR Pfam: PF00090; tsp_1; 1.
 DR SMART: SM00041; CT; 1.
 DR SMART: SM00214; IB; 1.
 DR SMART: SM00209; TSP1; 1.
 DR SMART: SM00214; WVC; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 DR PROSITE: PS00222; IGF_BINDING; 1.
 DR PROSITE: PS50092; TSP1; 1.
 DR NON TER 331
 FT SEQUENCE 331 AA; 36909 MW; D109CFDCA1DF549 CRC64;
 SO
 Query Match 29.6%; Score 626.5; DB 4; Length 331;
 Best Local Similarity 36.2%; Pred. No. 1.2e-52;
 Matches 122; Conservative 50; Mismatches 118; Indels 47; Gaps 8;
 QY 26 CPACHCLEAPKCAPGVGLVNDGCGCKVCAKQLENCSTQTCQDHTKGLECNFGASST 85
 DB 26 CHMCKCPQCKPRCPSPVSLVRDGGCKCAKQLENCSTQTCQDHTKGLECNFGASST 85
 QY 86 ALK-GICRAOSEBRCPEVNSRIYONSESOPNCHQCTIDGAVGCIPLCPQELSLPNLG 144
 DB 86 RYETGVCAKSVG--CEBNQVHYHNGVFOFNPFLSCLVSGAIGCTPLF----- 133
 QY 145 CPNRLKVVNGCCCEWNCDEDSIKDPMEDDGLGKELGDFASEVELTRNNELIAVGK 204
 DB 134 -----IPKLAGSHC-----SGAKGGRKS-DQNSCSL-----EPLDQLS 166
 QY 205 RSLKRLPFGMEPIILYNPLQGKCIQVOTWSGCKTGGTGSTRTVNTNPECRLYKET 264
 DB 167 TSYKTMAYRNLPPI-----WKKKLVQATKTPCSRTGCGISNRYTNENSNCEMRK 221
 QY 265 RICEVRPCGQPVYSLK--KGKCKSTKKSPEPVRYAGCLSVKKRYKYGSCVDGRG 322
 DB 222 RLCTIQPDNSILKTIKIPKCKTQPTFOLSKAEKVFSGCSSTQSTKPTFGICLDKRC 281
 QY 323 CTPOLTRTYKMRPCEDEGETFSKNVMIQSKCNYNC 359
 DB 282 CIPNKSMTITIQFDCPNBGSFKMKMLMTITSCVQRC 318
 RESULT 13
 095389 PRELIMINARY; PRT; 354 AA.
 AC 095389;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Connective tissue growth factor related protein WISP-3.
 GN WISP3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Bone marrow, and fetal kidney;


```
Db      115  --CAH-----AVG-----120
QY      175  QDGLGKELGFDASEVELTRNNELIavgKSHKRLPVEFGMEPRILYNPLQGRCIVQTT 234
Db      121  -----EVEAMHRN-----CIAVTS 134
QY      235  SWSGCKTGCTGSTRNTNDNPEGRVKEPRICEVRPGOPVYSSLKKGKCKSKTKKSPE 294
Db      135  PWSPCSTSCGLGVSTRISNVAOCWPEQESRLCMLRPCDDVDIHTLIKAGKCLAVYOPEA 194
QY      295  PVRETYAGCLSVKKYRPRKYGSCYDGRCTPOLTRTYKMRPRCEDEGETFSKNVMIOSCK 354
Db      195  SMNFTLAGCISTRSYOPKYGVCMDNRCCIPYKSKTIDVSRQCPDGLGFSRQVLMINACF 254
QY      355  CNYNCPHANEAPFPYRLFENDIHKFRD 381
Db      255  CNLSCRNPND-----IFADLESYPD 274
```

Search completed: August 5, 2003, 14:06:51
Job time : 34.0868 secs

THIS PAGE BLANK (USPTO)